目　录

第一章 测度论基础

近代概率论建立在测度论的基础上．测度论中的重要概念，测度、积分、各种收敛关系对应着概率论中的概率、数学期望、概率论中的收敛关系．掌握测度论的基本概念与方法，对学习和掌握近代概率论有很大帮助．反过来，不用测度论，那么近代概率论的某些表述就可能是不精确的．不学测度论，要掌握近代概率论是困难的．因此，从科学发展的观点看，在学习近代概率论之前，学习测度论是必要的．

在本书中，我们介绍可测函数、测度、积分、乘积测度、重积分等重要概念，阐述各种收敛关系及著名收敛定理．最后引入广义测度与 Radon-Nikodym 定理，为下面学习条件概率与鞅论铺平道路．

为了讲解可测函数，这里先从集及集的运算讲起．

1.1 集与集的运算

提要 本节介绍集的概念，集的运算规律，上限集，下限集，集的极限以及 De Morgan 定律．

集是一个不能精确定义，只能描述的基本概念．所谓集或集合，是指由一些对象组成的总体．组成集的对象称为集的元素．例如

(1) 骰子的 6 个面组成一个集，每个面是集的一个元素．

(2) n 次代数方程所有的根组成一个集，每个根是集的

一个元素。

(3) 设 $y(t)$, $u(t)$ 分别表示 t 时刻某系统的输出、输入,例如

$$y(t)+a_1y(t-1)+\cdots+a_ny(t-n)$$
$$=b_1u(t-1)+\cdots+b_mu(t-m)$$

表示所有线性模型组成一个集。这里不同的系数值 $(a_1,\cdots,a_n,b_1,\cdots,b_m)$,对应不同的线性模型。每一组系数值对应的模型是集的元素。

在本书中,我们用大写字母 A, B, C, $\cdots$ 等表示集。

如果集 A 的每一个元素都是集 B 的一个元素,则称 A 是 B 的子集,并记作

$$A\subset B \quad \text{或} \quad B\supset A$$

读作"B 包含 A"或"A 被 B 包含"。

例如,设 Q 是全体有理数的集;N 是全体自然数的集;R 是全体实数的集,那么

$$N\subset Q, \quad Q\subset R$$

或记作

$$N\subset Q\subset R$$

不含任何元素的集称为空集,恒用 $\varnothing$ 表示。我们规定空集是任何集的子集,即对任何 A, $\varnothing\subset A$。

在许多问题中,我们所考虑的集常常是某个给定集的子集。由问题中涉及的全部元素组成的集称为空间,记作 Ω,其元素记作 ω。

定义 1.1.1 设 A, B 是两个集,由 A 与 B 的所有元素合并在一起所构成的集,称为 A 与 B 的和集(或并集),简称为联或并,记作

$$A\cup B=\{\omega\mid\omega\in A \text{ 或 } \omega\in B\}$$

由 A 与 B 的所有公共元素所构成的集称为 A 与 B 的交集,简称为交,记作 $A\cap B$,即

$$A\cap B=\{\omega\mid\omega\in A \text{ 且 } \omega\in B\}=\{\omega:\omega\in A \text{ 且 } \omega\in B\}$$

上述表示式中，ω 表示元素，而竖线后给出了元素应满足的条件。

类似地可以定义任意个集的和集及交集。设

$$\{A_t\mid t\in T\}$$

是任意的一组集，其中 t 是指标，它在某个指标集 T 中变化，由一切 $A_t(t\in T)$ 的所有元素合并在一起所组成的集，称为这组集的和集，记作 $\bigcup\limits_{t\in T}A_t$，即

$$\bigcup_{t\in T}A_t=\{\omega\mid\omega \text{ 至少属于某一个 } A_t, t\in T\}$$

由同时属于每个集 $A_t(t\in T)$ 的所有元素组成的集，称为这组集的交集，记作 $\bigcap\limits_{t\in T}A_t$，即

$$\bigcap_{t\in T}A_t=\{\omega\mid\omega\in A_t,\ \text{对每个 } t \text{ 同时成立}\}$$

如果 $T=N$（自然数集），则上述的联与交分别称为可列联与可列交。

由定义易知，集的和与交运算具有下列性质：

(1) $A\cap B\subset A$，$A\cap B\subset B$；

(2) $A\subset A\cup B$，$B\subset A\cup B$；

(3) $A\cup\varnothing=A$，$A\cup\Omega=\Omega$；

(4) $A\cap\varnothing=\varnothing$，$A\cap\Omega=A$；

(5) 等幂律：$A\cup A=A$，$A\cap A=A$；

(6) 交换律：$A\cup B=B\cup A$，$A\cap B=B\cap A$；

(7) 分配律

$$A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$$

$$A\cup(B\cap C)=(A\cup B)\cap(A\cup C)$$

(8) 吸收律

$$A\cup(A\cap B)=A$$
$$A\cap(A\cup B)=A$$

(9) 结合律

$$A\cup(B\cup C)=(A\cup B)\cup C$$
$$A\cap(B\cap C)=(A\cap B)\cap C$$

定义 1.1.2 设 Ω 是一空间，$A\subset\Omega$，则 Ω 中不属于 A 的所有元素构成的集称为 A 的补集(简称补)，记作 A'，即

$$A'=\{\omega\mid\omega\in\Omega \text{ 且 } \omega\notin A\}$$

补运算具有下列性质：

(1) 互补性：$A\cup A'=\Omega$，$A\cap A'=\varnothing$，$\varnothing'=\Omega$，$\Omega'=\varnothing$；

(2) 对合律：$(A')'=A$；

(3) $A\subset B$ 的充要条件是 $A'\supset B'$。

在集运算中经常要用到反映集之并与交之间对偶关系的重要定理，称为 De Morgan 定律。现叙述如下：

定理 1.1.1 (De Morgan 定律) 设 A，B 是两个集，则

(1) $(A\cup B)'=A'\cap B'$；

(2) $(A\cap B)'=A'\cup B'$。

证明 (1) $\omega\in(A\cup B)'\Longleftrightarrow\omega\notin A\cup B\Longleftrightarrow\omega\notin A$ 且 $\omega\notin B\Longleftrightarrow\omega\in A'$ 且 $\omega\in B'\Longleftrightarrow\omega\in A'\cap B'$

所以

$$(A\cup B)'=A'\cap B'$$

(2) $\omega\in(A\cap B)'\Longleftrightarrow\omega\notin A\cap B\Longleftrightarrow\omega\notin A$ 或 $\omega\notin B\Longleftrightarrow\omega\in A'$ 或 $\omega\in B'\Longleftrightarrow\omega\in A'\cup B'$

所以

$$(A\cap B)'=A'\cup B'$$

例 设 $\Omega=\{1,2,3,4\}$，$A=\{1,3\}$，$B=\{2,4\}$，验证 De Morgan 定理。

解 因为 $(A\cup B)'=(\{1,3\}\cup\{2,4\})'=\varnothing$

$$A'=\{2,4\},\ B'=\{1,3\}$$

故

$$A'\cup B'=\Omega,\quad A'\cap B'=\varnothing$$

由此可知

$$(A\cup B)'=\Omega'=\varnothing=A'\cap B'$$
$$(A\cap B)'=\varnothing'=\Omega=A'\cup B'$$

推论 1.1.1 (De Morgan 定律) 设 $\{A_\alpha\}$ 是一组集，则

$$\left(\bigcup_\alpha A_\alpha\right)'=\bigcap_\alpha A_\alpha' \tag{1.1.1a}$$

$$\left(\bigcap_\alpha A_\alpha\right)'=\bigcup_\alpha A_\alpha' \tag{1.1.1b}$$

从式 (1.1.1a) 和 (1.1.1b) 可以看出，集的并与交之间存在对偶关系，即并之补等于补之交，交之补等于补之并。

定义 1.1.3 设 $A_1, A_2, \cdots, A_n, \cdots$ 是任意一列集。属于上述集列中无限多个集的那种元素全体组成一个集，称为 $A_1, A_2, \cdots, A_n, \cdots$ 的上限集，记作 $\limsup A_n$ 或 $\overline{\lim\limits_{n\to\infty}}A_n$，即

$$\limsup_{n\to\infty} A_n=\{\omega\,|\,\omega \text{ 属于无限多个 } A_n\}$$

例如，$A_{2n-1}=A$，$A_{2n}=B$，$n=1,2,\cdots$，那么 A 中的元素属于无限多个 A_k，$k=2n-1$，$n=1,2,\cdots$；B 中的元素属于无限多个 A_k，$k=2n$，$n=1,2,\cdots$。

所以

$$\limsup_{n\to\infty} A_n=A\cup B$$

定义 1.1.4 设 $A_1,A_2,\cdots,A_n,\cdots$ 是任意一列集，除去有限多个集之外，所有集 A_n 都含有的那种元素组成一个集，称为这一集列的下限集，记作 $\liminf A_n$ 或 $\underline{\lim\limits_{n\to\infty}}A_n$ 即

$$\liminf_{n\to\infty} A_n=$$

$\{\omega \mid$存在正整数 $n_0(\omega)$，使得当 $n > n_0(\omega)$时 $\omega \in A_n\}$

由于 $\omega \in \liminf\limits_{n\to\infty} A_n$ 时，$\omega \in A_n$ $n > n_0(\omega)$，故该元素属于无限多个 A_n，因此 $\omega \in \limsup\limits_{n\to\infty} A_n$。这就是说有

$$\liminf_{n\to\infty} A_n \subset \limsup_{n\to\infty} A_n$$

这一包含关系，反之不一定成立。

例如，$A_{2n-1} = A$，$A_{2n} = B$，$n = 1, 2, \cdots$，那么只有同时属于A和B的元素才有可能在 $n > n_0(\omega)$ 时，均属于A_n，因此可证

$$\liminf_{n\to\infty} A_n = A \cap B$$

下面我们来证明它的结论。设 $\omega \in A \cap B$，则 $\omega \in A$ 且 $\omega \in B$，于是 $\omega \in A_n$，$n = 1, 2, \cdots$。由下限集的定义知，$\omega \in \liminf\limits_{n\to\infty} A_n$，故有

$$A \cap B \subset \liminf_{n\to\infty} A_n \tag{1.1.2}$$

另一方面，$\omega \in (A \cap B)'$，即 $\omega \notin A \cap B$，不妨设 $\omega \notin A$，即$\omega \notin A_{2n-1}(n = 1, 2, \cdots)$，因而$\omega \notin \liminf\limits_{n\to\infty} A_n$，即$\omega \in (\liminf\limits_{n\to\infty} A_n)'$，于是有

$$(A \cap B)' \subset (\liminf_{n\to\infty} A_n)'$$

根据补运算的性质(3)

$$(A \cap B) \supset \liminf_{n\to\infty} A_n \tag{1.1.3}$$

由式(1.1.2)和式(1.1.3)可得

$$\liminf_{n\to\infty} A_n = A \cap B$$

定理 1.1.2 设 $\{A_n\}$ 是任一列集，那么

(1) $$\left(\liminf_{n\to\infty} A_n\right)' = \limsup_{n\to\infty} A'_n; \tag{1.1.4a}$$

(2) $$\left(\limsup_{n\to\infty} A_n\right)' = \liminf_{n\to\infty} A'_n. \tag{1.1.4b}$$

证明 (1) $(\liminf\limits_{n\to\infty} A_n)' = \{\omega|$ 有无限多个 A_n 不含 $\omega\}$

$$= \{\omega|\omega \text{ 属于无限多个 } A'_n\}$$

$$= \limsup_{n\to\infty} A'_n$$

(2) $(\limsup\limits_{n\to\infty} A_n)' = \{\omega|\omega$ 只属于有限多个 $A_n\}$

$= \{\omega|\omega$ 除有限多个 A'_n 外,属于无限多个 $A'_n\} = \liminf\limits_{n\to\infty} A'_n$ ∎

定理 1.1.3 设 $\{A_n\}$ 是任意一列集,那么

(1) $$\limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k; \tag{1.1.5a}$$

(2) $$\liminf_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k. \tag{1.1.5b}$$

证明 (1) $\omega \in \limsup\limits_{n\to\infty} A_n \Longleftrightarrow$ 对于所有的 n,存在 $k \geq n$,使 $\omega \in A_k \Longleftrightarrow$

对于所有 n,$\omega \in \bigcup\limits_{k=n}^{\infty} A_k \Longleftrightarrow \omega \in \bigcap\limits_{n=1}^{\infty} \bigcup\limits_{k=n}^{\infty} A_k$

所以

$$\limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$$

(2) $$\liminf_{n\to\infty} A_n = [(\liminf_{n\to\infty} A_n)']' = \left(\limsup_{n\to\infty} A'_n\right)'$$

$$= \left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A'_k\right)' = \bigcup_{n=1}^{\infty} \left(\bigcup_{k=n}^{\infty} A'_k\right)'$$

$$= \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k$$

定义 1.1.5 如果集列 $\{A_n\}$ 的上限集和下限集相等,即

$$\limsup_{n\to\infty} A_n = \liminf_{n\to\infty} A_n$$

则称集列 $\{A_n\}$ 收敛，这时称 $A = \limsup\limits_{n\to\infty} A_n = \liminf\limits_{n\to\infty} A_n$ 为集列 $\{A_n\}$ 的极限(集)，记作

$$A = \liminf_{n\to\infty} A_n$$

设 $\{A_n\}$ 是一集列，如果 $A_n \subset A_{n+1}(n=1,2,\cdots)$ 则称 $\{A_n\}$ 是单调增加集列，并记作 $A_n\uparrow$；如果 $A_n \supset A_{n+1}(n=1,2,\cdots)$ 则称 $\{A_n\}$ 是单调减少集列，并记作 $A_n\downarrow$，单调增加集列和单调减少集列统称为单调集列。

下面单调集列的极限定理是今后很有用的工具。

定理 1.1.4 单调集列都是收敛的，且如果 $A_n\uparrow$，则

$$\lim_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} A_n \tag{1.1.6a}$$

如果 $A_n\downarrow$，则

$$\lim_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} A_n \tag{1.1.6b}$$

证明 若 $A_n\uparrow$，即 $A_n \subset A_{n+1}$，$n=1,2,\cdots$，由定理 1.1.3 知

$$\limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k$$

因为 $A_n\uparrow$，故有

$$\bigcup_{k=1}^{\infty} A_k = \bigcup_{k=2}^{\infty} A_k \cup A_1 = \bigcup_{k=3}^{\infty} A_k = \cdots$$

所以

$$\limsup_{n\to\infty} A_n = \bigcup_{k=1}^{\infty} A_k$$

另一方面，由定理 1.1.3 知

$$\liminf_{n\to\infty} A_n = \bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty} A_k$$

因为 $A_n\uparrow$，故有

$$\bigcap_{k=1}^{\infty} A_k = A_1,\ \bigcap_{k=2}^{\infty} A_k = A_2,\cdots,\ \bigcap_{k=n}^{\infty} A_k = A_n$$

所以

$$\liminf_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} A_n$$

故

$$\lim_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} A_n$$

对于 $A_n\downarrow$ 的情形，可类似证明。

习 题 1.1

1. 证明

(1) $\bigcup_{t\in T} A_t - A = \bigcup_{t\in T}(A_t - A)$,

(2) $\bigcap_{t\in T} A_t - A = \bigcap_{t\in T}(A_t - A)$,

(3) $A - \bigcup_{t\in T} A_t = \bigcap_{t\in T}(A - A_t)$,

(4) $A - \bigcap_{t\in T} A_t = \bigcup_{t\in T}(A - A_t)$.

2. 设 $\{A_n\}$ 是单调减少集列，证明

$$A_1 = \sum_{n=1}^{\infty}(A_n - A_{n+1}) + \bigcap_{n=1}^{\infty} A_n$$

这里$\sum$表示不相交集之并.

3. 设 $\{A_k\}$ 是任意集列，证明

(1) $\lim_{n\to\infty}\bigcup_{k=1}^{n} A_k = \bigcup_{k=1}^{\infty} A_k$,

(2) $\lim_{n\to\infty}\bigcap_{k=1}^{n} A_k = \bigcap_{k=1}^{\infty} A_k$.

4. 设 $A_{3n} = A$, $A_{3n-1} = B$, $A_{3n-2} = C$, $n = 1, 2, \cdots$, 求集列 $\{A_n\}$ 的上限集和下限集.

5. 设 $A_{2n-1} = \left(0, \frac{1}{n}\right)$, $A_{2n} = (0, n)$, $n = 1, 2, \cdots$, 求集列 $\{A_n\}$ 的上限集与下限集.

6. 证明关系式

$$(E - F) \cup F = E$$

成立的充分必要条件为 $F \subset E$.

7. 设 $E \supset F \supset G$, 证明 $E - G = (E - F) \cup (F - G)$.

8. 设 $\{E_n\}$, $n = 1, 2, \cdots$ 是两两不相交集序列，则 $\lim_{n\to\infty} E_n = \varnothing$.

9. 设 $\lim_{n\to\infty} A_n$ 存在，B 是任意一个集，证明

(1) $\lim_{n\to\infty}(A_n - B) = \lim_{n\to\infty} A_n - B$,

(2) $\lim_{n\to\infty}(B - A_n) = B - \lim_{n\to\infty} A_n$.

1.2 可测函数

提要 本节阐述 σ 代数，Borel 代数，集的示性函数，实可测函数，复可测函数等基本概念，并举例佐证. 本节还证明了任一非负可测函数是一单调增简单函数列的极限(可测函数的结构性质).

在上节建立了集与集运算概念之后，在本节中我们就可以建立 σ 代数的概念，从而建立可测函数的概念. 可测函数是测度论中积分理论讨论的对象.

从积分理论看，Riemann 积分本质上只考虑连续函数

的积分，或者说处理连续函数以及有有限个间断点或至多有可列个间断点的函数的积分．然而，这样的积分对于近代概率论来说是不够的．原来 Lebesgue 考虑了定义在实数空间 R 上的可测函数积分．这是实变函数论中讨论的问题．在本书中，我们将考虑抽象空间 Ω 上实可测函数的积分和极限定理．

在本书的讨论中，Ω 可以表示不同的集合．例如，Ω 可以是 $[0,1]$ 闭区间中的全体实数，Ω 可以是全体自然数，Ω 也可以是整个实轴或整个实平面或其他抽象集合．由于本书讨论的抽象积分(简称积分)与集 Ω 的性质无关，因此我们对 Ω 不作任何假定．

为了讨论可测函数，我们要研究 Ω 的子集所成的集合．我们把 Ω 的子集所成的类称为集类或集族．设 $\mathscr{F}$ 是 Ω 子集所成的集类，如果空集 $\varnothing$ 和整个空间都属于 $\mathscr{F}$，且关于补运算、可列并运算封闭，则 $\mathscr{F}$ 称为 σ 代数。严格的定义如下：

定义 1.2.1 设 $\mathscr{F}$ 是 Ω 子集所成的集类．如果

(1) Ω 属于 $\mathscr{F}$，即 $\Omega\in\mathscr{F}$；

(2) 如果 $A\in\mathscr{F}$，那么 A 的补 $A'\in\mathscr{F}$；

(3) 如果 $A_n\in\mathscr{F}$，$\forall n=1,2,\cdots$，那么 A_n 的并 $\bigcup\limits_{n=1}^{\infty}A_n\in\mathscr{F}$（这里符号"$\forall$"读作"对于"），我们称 $\mathscr{F}$ 是 σ 代数。

包含有 Ω 和 σ 代数 $\mathscr{F}$ 的有序对 $(\Omega,\mathscr{F})$ 称为可测空间。$\mathscr{F}$ 中的任一集合，称为可测集。

若 $A_n\in\mathscr{F}$，$n=1,2,\cdots$，$\mathscr{F}$ 是 σ 代数，那么

$$\left(\bigcup_{n=1}^{\infty}A_n\right)'\in\mathscr{F}，即 \bigcap_{n=1}^{\infty}A'_n\in\mathscr{F}$$

由 De Morgan 定律

$$\left(\bigcup_{n=1}^{\infty} A_n\right)' = \bigcap_{n=1}^{\infty} A'_n \in \mathscr{F} \tag{1.2.1}$$

可知：若 $A_n \in \mathscr{F}$，$\mathscr{F}$ 是 σ 代数，那么

$$\bigcap_{n=1}^{\infty} A_n \in \mathscr{F}$$

例 下面给出一些 σ 代数的例子.

(1) 由 $\varnothing$，Ω 两个集组成的集类是 σ 代数. 这是因为 $\varnothing$，Ω 的补与可列并运算结果是 $\varnothing$ 和 Ω.

(2) 设 $\Omega=\{1,2,3,\cdots\}$ 是自然数集，$\mathscr{F}$ 包含如下子集：

$$\varnothing,\ \{1,3,5,\cdots\},\ \{2,4,6,\cdots\},\ \Omega$$

则 $\mathscr{F}$ 是 σ 代数.

(3) Ω 是任一集合，$\mathscr{F}$ 包含 Ω 的所有子集，则 $\mathscr{F}$ 是 σ 代数.

(4) Ω 是一个不可列集，$\mathscr{F}$ 由一切可列子集及其补集是可列集的子集组组成，则 $\mathscr{F}$ 是 σ 代数.

(5) 设 $\mathscr{F}_1$ 和 $\mathscr{F}_2$ 是 Ω 子集组成的两个 σ 代数

$$\mathscr{F}_3 = \mathscr{F}_1 \cap \mathscr{F}_2$$

则 $\mathscr{F}_3$ 包含了 $\mathscr{F}_1$ 和 $\mathscr{F}_2$ 的一切公共元，容易验证 $\mathscr{F}_3$ 也是 σ 代数.

(6) 设 $\mathscr{A}$ 是 Ω 的非空子集，$\mathscr{F}_\alpha$ 是任一包含 $\mathscr{A}$ 的 σ 代数，那么 $\bigcap_{\alpha} \mathscr{F}_\alpha$ 称为包含 $\mathscr{A}$ 的最小 σ 代数，有时也称为由 $\mathscr{A}$ 生成的 σ 代数.

(7) 设 Ω 是实轴 R（实数全体）. $\mathscr{B}$ 是由 R 中一切开区间 (a, b) 生成的 σ 代数，称为 Borel 代数. 显然 Borel 代数也可看作 R 中一切闭区间 $[a, b]$ 生成的 σ 代数. $\mathscr{B}$ 中的任何集合称为 Borel 集.

(8) 设 Ω 是广义实数集 $\bar{R}$. 如果 E 是 R 的一个 Borel 子集，令

$E_1 = E\cup\{-\infty\}$，$E_2 = E\cup\{+\infty\}$，$E_3 = E\cup\{-\infty,+\infty\}$
$\bar{\mathscr{B}}$ 是当 E 取遍$\mathscr{B}$中所有集时，E，E_1，E_2，E_3 组成的集类，显然$\bar{\mathscr{B}}$也是 σ 代数，称为广义 Borel 代数。

设 $(\Omega,\mathscr{F})$ 是某一个给定的可测空间. 下面的讨论都在这一假定下进行.

定义 1.2.2 f 是 Ω 到 $\bar{R}$ 的一个函数，如果对于任何有限实数 α 有

$$\{\omega\in\Omega \mid f(\omega)>\alpha\} \tag{1.2.2}$$

属于 $\mathscr{F}$，则称 f 为可测函数。

下面的引理告诉我们在函数可测性定义中集合(1.2.2)的形式是可以改变的.

引理 1.2.1 设 f 是 Ω 到 $\bar{R}$ 的一个函数，下列一些命题是等价的：

(1) 对任何 $\alpha\in R$，集 $A_\alpha=\{\omega\in\Omega \mid f(\omega)>\alpha\}\in\mathscr{F}$；
(2) 对任何 $\alpha\in R$，集 $B_\alpha=\{\omega\in\Omega \mid f(\omega)\leqslant\alpha\}\in\mathscr{F}$；
(3) 对任何 $\alpha\in R$，集 $C_\alpha=\{\omega\in\Omega \mid f(\omega)\geqslant\alpha\}\in\mathscr{F}$；
(4) 对任何 $\alpha\in R$，集 $D_\alpha=\{\omega\in\Omega \mid f(\omega)<\alpha\}\in\mathscr{F}$.

证明 因为 B_α 和 A_α 互为补集，因此(1)与(2)是等价命题. 类似地，(3)和(4)也是等价命题. 如果命题(1)成立，那么对于任何 n，$A_{\alpha-\frac{1}{n}}\in\mathscr{F}$ 因为

$$C_\alpha=\bigcap_{n=1}^{\infty}A_{\alpha-\frac{1}{n}}$$

根据$\mathscr{F}$关于可列交运算封闭，$C_\alpha\in\mathscr{F}$. 因此命题(1)成立可推出命题(3)成立. 又因为

$$A_\alpha=\bigcup_{n=1}^{\infty}C_{\alpha+\frac{1}{n}}$$

因此命题(3)成立可推出命题(1)成立. 于是命题(1)与命题

(3)等价.

例

(1) 任何常数函数是可测的. 例如

$$f(\omega)=C\ (\text{常数}),\ \forall\omega\in\Omega$$

若 $\alpha\geqslant C$，那么

$$\{\omega\in\Omega\mid f(\omega)>\alpha\}=\varnothing$$

若 $\alpha<C$，那么

$$\{\omega\in\Omega\mid f(\omega)>\alpha\}=\Omega$$

(2) 如果 $E\in\mathscr{F}$，集合E的示性函数χ_E定义为

$$\chi_E(\omega)=\begin{cases}1, & \forall\omega\in E\\ 0, & \forall\omega\notin E\end{cases}$$

那么 $\chi_E(\omega)$ 是可测函数. 事实上 $\{\omega\in\Omega\mid\chi_E(\omega)>\alpha\}$ 必为Ω, E 和$\varnothing$三者之一.

(3) 设 $\Omega=R$ 是全体实数组成的集合. $\mathscr{F}=\mathscr{B}$ 是 Borel 代数,那么R到R的任何连续函数f是 Borel 可测的. 事实上,如果f是连续函数,那么$\{x\in R: f(x)>\alpha\}$是R中的开集,因此它是可列个开区间的联，于是 $\{x\in R\mid f(x)>\alpha\}\in\mathscr{F}$.

(4) 如果 $\Omega=R$, $\mathscr{F}=\mathscr{B}$, 那么任何单调函数是 Borel 可测的. 为了证明这一点,假定f是单调增函数,即由 $x\leqslant x'$ 可推出 $f(x)\leqslant f(x')$. $\{x\in R\mid f(x)>\alpha\}$ 是 $(a,+\infty)$ 或 $[\alpha,+\infty)$的无穷区间.(试证明这两种情形都可能出现)

下面的引理说明了可测函数的某些代数运算的结果仍然是可测函数.

引理 1.2.2 设f和g都是实可测函数(即Ω到R的函数),c是任一实数,那么

$$cf,\quad f^2,\quad f+g,\quad fg,\quad |f|$$

都是可测函数.

证明 (1) 如果 $c=0$，命题显然成立.

如果 $c>0$，那么

$\{\omega\in\Omega\mid cf(\omega)>\alpha\}=\{\omega\in\Omega\mid f(\omega)>\alpha/c\}\in\mathscr{F}$

$c<0$ 的证明类似.

(2) 如果 $\alpha<0$，那么 $\{\omega\in\Omega\mid f^2(\omega)>\alpha\}=\Omega$；如果 $\alpha\geqslant 0$ 那么

$$\{\omega\in\Omega\mid f^2(\omega)>\alpha\}=\{\omega\in\Omega\mid f(\omega)>\sqrt{\alpha}\}\cup\{\omega\in\Omega\mid f(\omega)<-\sqrt{\alpha}\}\in\mathscr{F}$$

(3) 设 Q 是全体有理数的集合，$r\in Q$ 那么

$$S_r=\{\omega\in\Omega\mid f(\omega)>r\}\cap\{\omega\in\Omega\mid g(\omega)>\alpha-r\}\in\mathscr{F}$$

因为

$$\{\omega\in\Omega\mid f(\omega)+g(\omega)>\alpha\}=\bigcup_{r\in Q}S_r\in\mathscr{F}$$

因此 $f+g$ 是可测函数.

(4) 因为

$$fg=\frac{1}{4}[(f+g)^2-(f-g)^2]$$

由上面的(1),(2)和(3)可知 fg 可测.

(5) 如果 $\alpha<0$，$\{\omega\in\Omega\mid |f(\omega)|>\alpha\}=\Omega$；如果 $\alpha\geqslant 0$，$\{\omega\in\Omega\mid |f(\omega)|>\alpha\}=\{\omega\in\Omega\mid f(\omega)>\alpha\}\cup\{\omega\in\Omega\mid f(\omega)<-\alpha\}\in\mathscr{F}$. 于是 $|f(\omega)|$ 是可测函数. ▌

可测函数的正部与负部定义为

$$f^+(\omega)=\sup\{f(\omega),0\}\tag{1.2.3a}$$

$$f^-(\omega)=\sup\{-f(\omega),0\}\tag{1.2.3b}$$

这里 f^+ 称为 f 的正部；f^- 称为 f 的负部.显然有

$$f=f^+-f^-\tag{1.2.4}$$

$$|f| = f^+ + f^- \tag{1.2.5}$$

利用下面的等式:

$$f^+ = \frac{1}{2}(|f| + f), \quad f^- = \frac{1}{2}(|f| - f) \tag{1.2.6}$$

以及引理 1.2.2. 可知: 如果 f 是可测函数, 那么 f^+ 和 f^- 都是可测函数. 反之亦然.

为了使可测函数列的极限运算、上确界运算方便,下面我们引入广义实函数的可测性.

定义 1.2.3 设 f 是 Ω 上的广义实函数, 若对任何实数 α 有

$\{\omega \in \Omega \mid f(\omega) > \alpha\} \in \mathscr{F}$, α 允许取 $\pm\infty$

则称 f 是广义实可测函数,或简称 f 是可测的.

$(\Omega, \mathscr{F})$ 上的全体广义实可测函数记作 $M(\Omega, \mathscr{F})$.

如果 $f \in M(\Omega, \mathscr{F})$, 由于

$$\{\omega \in \Omega \mid f(\omega) = +\infty\} = \bigcap_{n=1}^{\infty} \{\omega \in \Omega \mid f(\omega) > n\}$$

$$\{\omega \in \Omega \mid f(\omega) = -\infty\} = \left[\bigcup_{n=1}^{\infty} \{\omega \in \Omega \mid f(\omega) > -n\}\right]'$$

所以这两个集均属于 $\mathscr{F}$.

下面的引理对于讨论广义实函数是有用的.

引理 1.2.3 广义实函数 f 可测的充分必要条件为 $A = \{\omega \in \Omega \mid f(\omega) = +\infty\}$, $B = \{\omega \in \Omega \mid f(\omega) = -\infty\}$, $A, B \in \mathscr{F}$, $f_1(\omega)$ 定义为

$$f_1(\omega) = \begin{cases} f(\omega), & \forall \omega \notin A \cup B \\ 0, & \forall \omega \in A \cup B \end{cases}$$

$f_1(\omega)$ 是可测函数.

证明 若 $f \in M(\Omega, \mathscr{F})$, 上面已经证明 A, B 均属于 $\mathscr{F}$. 设 $\alpha \in R$ 且 $\alpha \geqslant 0$, 那么

$$\{\omega\in\Omega \mid f_1(x)>\alpha\}=\{\omega\in\Omega \mid f(\omega)>\alpha\}-A$$

若 $\alpha<0$，那么

$$\{\omega\in\Omega \mid f_1(x)>\alpha\}=\{\omega\in\Omega \mid f(\omega)>\alpha\}\cup B$$

因此 f_1 可测.

反之，若 $A,B\in\mathscr{F}$ 且 $f_1(\omega)$ 可测，那么当 $\alpha\geqslant 0$ 时，

$$\{\omega\in\Omega \mid f(\omega)>\alpha\}=\{\omega\in\Omega \mid f_1(\omega)>\alpha\}\cup A\in\mathscr{F}$$

当 $\alpha<0$ 时

$$\{\omega\in\Omega \mid f(\omega)>\alpha\}=\{\omega\in\Omega \mid f_1(\omega)>\alpha\}-B\in\mathscr{F}$$

因此 f 可测.

推论 1.2.1（引理 1.2.2 和引理 1.2.3 的推论） 如果 $f\in M(\Omega,\mathscr{F})$，$c\in R$，那么

$$cf,\ f^2,\ |f|,\ f^+,\ f^-$$

均属于 $M(\Omega,\mathscr{F})$.

这里需要指出，我们约定 $0(\pm\infty)=0$，所以当 $c=0$ 时 $cf=0$. 如果 $f,g\in M(\Omega,\mathscr{F})$ 令

$$E_1=\{\omega\in\Omega \mid f(\omega)=-\infty\ \text{且}\ g(\omega)=+\infty\}$$

$$E_2=\{\omega\in\Omega \mid f(\omega)=+\infty\ \text{且}\ g(\omega)=-\infty\}$$

这时 $(f+g)(\omega)=f(\omega)+g(\omega)$ 在 $E_1\cup E_2$ 上没有定义.

然而，如果约定 $f+g$ 在 $E_1\cup E_2$ 上等于零，那么 $f+g$ 就是可测函数. 利用下面的引理，我们可以讨论 fg 的可测性.

引理 1.2.4 设 $\{f_n\}$ 是 $M(\Omega,\mathscr{F})$ 中的序列，定义

$$f(\omega)=\inf f_n(\omega),\quad F(\omega)=\sup f_n(\omega)$$

$$f^*(\omega)=\liminf f_n(\omega),\quad F^*(\omega)=\limsup f_n(\omega),$$

那么 f, F, f^* 和 F^* 均属于 $M(\Omega,\mathscr{F})$.

证明 考虑下列集合:

$$\{\omega\in\Omega\mid f(\omega)\geqslant\alpha\}=\bigcap_{n=1}^{\infty}\{\omega\in\Omega\mid f_n(\omega)\geqslant\alpha\}\in\mathscr{F}$$

$$\{\omega\in\Omega\mid F(\omega)>\alpha\}=\bigcup_{n=1}^{\infty}\{\omega\in\Omega\mid f_n(\omega)>\alpha\}\in\mathscr{F}$$

因此当 $f_n\in M(\Omega,\mathscr{F})$ 时，f 和 F 可测.

又因为

$$f^*(\omega)=\sup_{n\geqslant 1}\{\inf_{m\geqslant n} f_m(\omega)\}$$

$$F^*(\omega)=\inf_{n\geqslant 1}\{\sup_{m\geqslant n} f_m(\omega)\}$$

因此 f^* 和 F^* 也可测. ▮

推论 1.2.2 如果 $\{f_n\}$ 是 $M(\Omega,\mathscr{F})$ 中的函数列，

$$\lim_n f_n=f$$

那么 $f\in M(\Omega,\mathscr{F})$.

证明 在这种情形下

$$f(\omega)=\lim_n f_n(\omega)=\lim\inf f_n(\omega)\in M(\Omega,\mathscr{F}) \quad ▮$$

现在我们回过来讨论，当 $f, g\in M(\Omega,\mathscr{F})$ 时，$fg\in M(\Omega,\mathscr{F})$.

设 $N=\{1,2,\cdots\}$，$n\in N$，f 的截尾函数定义为

$$f_n(\omega)=\begin{cases} f(\omega), & \forall |f(\omega)|\leqslant n \\ n, & \forall f(\omega)>n \\ -n, & \forall f(\omega)<-n \end{cases}$$

同样

$$g_m(\omega)=\begin{cases} g(\omega), & \forall |g(\omega)|\leqslant m \\ m, & \forall g(\omega)>m \\ -m, & \forall g(\omega)<-m \end{cases}$$

容易看出 f_n 和 g_m 均为可测函数[习题 1.2, 11]. 由引理 1.2.2 可知，$f_n g_m$ 是可测的. 因为

$$f(\omega)g_m(\omega)=\lim_n f_n(\omega)g_m(\omega),\ \forall\omega\in\Omega$$

根据推论 1.2.2 可知，$fg_m\in M(\Omega,\mathscr{F})$.

又因为

$$(fg)(\omega)=f(\omega)g(\omega)=\lim_m f(\omega)g_m(\omega),\ \forall\omega\in\Omega$$

再次利用推论 1.2.2，可知 $fg\in M(\Omega,\mathscr{F})$.

现在我们已经证明了 $M(\Omega,\mathscr{F})$ 中函数列的极限属于 $M(\Omega,\mathscr{F})$. 下面我们将证明 $M(\Omega,\mathscr{F})$ 中的任一非负可测函数必定是 $M(\Omega,\mathscr{F})$ 中简单函数列 $\{\varphi_n(\omega)\}$ 的极限且 $\{\varphi_n(\omega)\}$ 是单调增函数序列.

引理 1.2.5 如果 $f\in M(\Omega,\mathscr{F})$ 且 $f\geqslant 0$，那么在 $M(\Omega,\mathscr{F})$ 中必存在一个函数列 $\{\varphi_n\}$ 满足：

(1) $0\leqslant\varphi_n(\omega)\leqslant\varphi_{n+1}(\omega),\ \forall\omega\in\Omega,\ n\in N$.

(2) $f(\omega)=\lim\varphi_n(\omega),\ \forall\omega\in\Omega$.

(3) $\varphi_n(\omega)$ 只取有限个实数值 [此时，$\varphi_n(\omega)$ 称为简单函数].

证明 设 n 是某给定的自然数. 若 $k=0,1,\cdots,n2^n-1$ 时，令

$$E_{kn}=\{\omega\in\Omega\,|\,k2^{-n}\leqslant f(\omega)<(k+1)2^{-n}\}$$

当 $k=n2^n$ 时，令

$$E_{kn}=\{\omega\in\Omega\,|\,f(\omega)\geqslant n\}$$

由此可见，$\{E_{kn}\,|\,k=0,1,\cdots,n2^n\}$ 全是互不相交的集合，这些集合均属于 $\mathscr{F}$，且

$$\Omega=\bigcup_{n=0}^{n2^n}E_{kn}$$

如果我们定义 Ω 上的函数 $\varphi_n(\omega)$，使得它在 E_{kn} 上等于 $k2^{-n}$，那么 $\varphi_n\in M(\Omega,\mathscr{F})$. 容易验证这样定义出来的 $\{\varphi_n\}$ 满足引理中的(1),(2),(3)条件. ▌

复可测函数在概率论中扮演着重要的角色．如果 f 是定义在 Ω 上的复值函数，那么一定存在两个唯一确定的实值函数 f_1，f_2，使得

$$f=f_1+if_2$$

显然这里 $f_1(\omega)=\mathrm{Re}f(\omega)$，$f_2(\omega)=\mathrm{Im}f(\omega)$，$\forall\omega\in\Omega$．$f_1(\omega)$ 称为 $f(\omega)$ 的实部；$f_2(\omega)$ 称为 $f(\omega)$ 的虚部．当且仅当 $f(\omega)$ 的实部 $f_1(\omega)$ 和虚部 $f_2(\omega)$ 均可测时，$f(\omega)$ 可测．容易验证复可测函数的和，乘积，极限也是可测的．

在某些场合，人们需要定义从一个可测空间 $(\Omega,\mathscr{F})$ 到另一个可测空间 $(\Lambda,\mathscr{G})$ 的函数 f．在这种情况下，f 可测表示

$$f^{-1}(E)=\{\omega\in\Omega\mid f(\omega)\in E\}\in\mathscr{F},\ \forall E\in\mathscr{G}$$

从表面上看，这里的定义与定义 1.2.2 不同．然而不难证明[见习题 1.2，11]，当 $\Lambda=R$，$\mathscr{G}=\mathscr{B}$ 时，这里的定义与定义 1.2.2 等价．

习　　题　　1.2

1．证明

$$[a,b]=\bigcap_{n=1}^{\infty}\left(a-\frac{1}{n},\ b+\frac{1}{n}\right)$$

这样 R 上任一包含了所有开区间的 σ 代数必包含所有闭区间．

同样证明

$$(a,b)=\bigcup_{n=1}^{\infty}\left[a+\frac{1}{n},\ b-\frac{1}{n}\right]$$

这样 R 上任一包含所有闭区间的 σ 代数必包含所有的开区间．

2．证明 Borel 代数是包含一切半开区间 $[a,b)=\{x\in R:\ a\leqslant x<b\}$ 的 σ 代数．并证明 Borel 代数也是包含一切半射线 $\{x\in R:\ x>a\}$ 的 σ 代数．

3．设 $\{A_n\}$ 是 Ω 的一个子集序列。令 $E_0=\varnothing$，

$$E_n=\bigcup_{k=1}^{n}A_k,\quad F_n=A_n-E_{n-1}$$

证明

(1) $\{E_n\}$ 是一单调递增集合序列。

(2) $\{F_n\}$ 两两不相交，即 $F_n\cap F_m=\varnothing$，$\forall n\neq m$。

(3) $\bigcup\limits_{n=1}^{\infty}E_n=\bigcup\limits_{n=1}^{\infty}F_n=\bigcup\limits_{n=1}^{\infty}A_n$。

4．举出一个函数 f 本身不可测，但 $|f|$ 和 f^2 可测的例子。

5．设 a,b,c 为三个实数，$\mathrm{mid}(a,b,c)$ 表示值在中间的那个数，证明

$$\mathrm{mid}(a,b,c)=\inf\{\sup(a,b),\sup(a,c),\sup(b,c)\}$$

若 f_1,f_2,f_3 是 Ω 到 R 的三个可测函数，g 定义为

$$g(\omega)=\mathrm{mid}(f_1(\omega),f_2(\omega),f_3(\omega))$$

证明 $g(\omega)$ 可测。

6．若 f 是可测函数，$A>0$，f_A 表示 f 的截尾函数并定义为

$$f_A(x)=\begin{cases}f(x), & \text{当 } |f(x)|\leqslant A\\ A, & \text{当 } f(x)>A\\ -A, & \text{当 } f(x)<-A\end{cases}$$

证明 $f_A(x)$ 可测。

7．设 f 是 Ω 上有界非负可测函数，即存在常数 $K>0$，使得

$$0\leqslant f(\omega)\leqslant K,\ \forall\omega\in\Omega$$

试证明引理 1.2.5 中构造的简单函数列 $\{\varphi_n\}$ 在 Ω 上均匀收敛于 f。

8．设 f 是 X 到 Y 的函数。若 $E\subset Y$，令

$$f^{-1}(E)=\{x\in X: f(x)\in E\}$$

证明 $f^{-1}(\varnothing)=\varnothing$，$f^{-1}(Y)=X$。

若 E 和 F 都是 Y 的子集，证明

$$f^{-1}(E-F)=f^{-1}(E)-f^{-1}(F)$$

若 $\{E_\alpha\}$ 是 Y 子集的非空类，证明

$$f^{-1}\Big(\bigcup_\alpha E_\alpha\Big)=\bigcup_\alpha f^{-1}(E_\alpha),\quad f^{-1}\Big(\bigcap_\alpha E_\alpha\Big)=\bigcap_\alpha f^{-1}(E_\alpha)$$

特别是，当 $\mathscr{G}$ 是 Y 子集所成的 σ 代数时，证明 $\{f^{-1}(E):\ E\in\mathscr{G}\}$ 是 Ω 子集所成的 σ 代数。

9．f 是定义在 X 上并取值于 Y 的函数，$(X,\mathscr{F})$ 是可测空间，令

$$\mathscr{G}=\{E\subset Y:\ f^{-1}(E)\in\mathscr{F}\}$$

证明 $\mathscr{G}$ 是 σ 代数。

10．设 $(\Omega,\mathscr{F})$ 是可测空间，f 是 Ω 到 Y 的函数，A 是 Y 的子集所成的类

$$A=\{E\subset Y:\ f^{-1}(E)\in\mathscr{F}\}$$

试证明

$f^{-1}(F)\in\mathscr{F}$，当 $F\in$ 由 A 生成的 σ 代数(提示：运用习题 1.2，9)

11．$(\Omega,\mathscr{F})$ 是一可测空间，f 是定义在 Ω 上的实值函数．证明 f 为可测函数的充分必要条件是对于每一个 Borel 集 E，$f^{-1}(E)\in\mathscr{F}$．

12．设 $(\Omega,\mathscr{F})$ 是一可测空间，f 是 Ω 到 R 的可测函数，φ 是 R 到 R 的连续函数．试证明复合函数

$$(\varphi\circ f)(\omega)=\varphi[f(\omega)]$$

是可测函数．[提示：若 φ 连续，则对于每一个 $E\in\mathscr{B}$ 有 $\varphi^{-1}(E)\in\mathscr{B}$]

13．设 f 是习题 1.2，7 中定义的函数，ϕ 是 Borel 可测函数，证明 $\phi\circ f$ 可测。

14．设 f 是定义在可测空间 $(\Omega,\mathscr{F})$ 上的复值函数．试证明 f 可测的充分必要条件是：对于所有实数 a,b,c,d，有

$$\{\omega\in\Omega:\ a<\operatorname{Re}f(\omega)<b,\ c<\operatorname{Im}f(\omega)<d\}\in\mathscr{F}$$

更一般地证明 f 可测的充分必要条件是：对于复平面中的任何开集 G，有

$$f^{-1}(G)\in\mathscr{F}$$

15．试证明复值可测函数的和，乘积，极限是可测函数。

16．$\mathscr{M}$ 是 Ω 子集的非空类，若对于 $\mathscr{M}$ 中的每一个单调增序列 $\{E_n\}$ 和单调减序列 $\{F_n\}$ 有

$$\bigcup_{n=1}^{\infty}E_n\in\mathscr{M},\quad \bigcap_{n=1}^{\infty}F_n\in\mathscr{M}$$

则称 $\mathscr{M}$ 为单调类。

试证明 σ 代数是单调类．设 $\mathscr{A}$ 是一非空类，那么必存在包含 $\mathscr{A}$ 的

最小单调类(称为由$\mathscr{A}$生成的单调类)。

1.3 测　　度

提要　本节阐述测度和测度空间等概念，并证明测度的单调性和连续性。

我们已经引进了可测空间的概念。$\mathscr{F}$ 表示Ω子集所成的σ代数。现在我们讨论在 $\mathscr{F}$ 上定义的实值或广义实值集函数。这种集函数将称为"测度"，它是长度、面积、质量等的抽象概括。因此它应该对于空集等于零，对于$\mathscr{F}$中不相交的集合具有可列可加性。

定义 1.3.1　μ是定义在 $\mathscr{F}$ 上的广义实值集函数，若满足

(1) $\mu(\varnothing)=0$;

(2) $\mu(E)\geqslant 0$, $\forall E\in\mathscr{F}$;

(3) 若 $\{E_n\}$ 是不相交序列，即 $E_n\cap E_m=\varnothing$, $\forall n\neq m$，那么

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right)=\sum_{n=1}^{\infty}\mu(E_n) \tag{1.3.1}$$

μ称为测度。

这里性质(3)称谓可列可加性。如果任何 $A\in\mathscr{F}$ 都具有有限测度(不取$+\infty$)，则称测度μ是有限的；如果 $\mu(\Omega)<+\infty$，则称μ是全有限的；如果 $\mu(\Omega)=1$，则称μ为概率测度。

设 $A\in\mathscr{F}$，如果存在一列集 $A_n\in\mathscr{F}(n=1,2,\cdots)$，使得

$$A \subset \bigcup_{n=1}^{\infty} A_n \text{ 且 } \mu(A_n) < +\infty$$

则称A的测度是σ有限的。 如果每个 $A \in \mathscr{F}$ 的测度都是σ有限，则称μ是σ有限的。

例

(1) 设Ω是一非空集，$\mathscr{F}$是Ω子集所成的σ代数。

μ_1在$\mathscr{F}$上定义为

$$\mu_1(E) = 0, \quad \forall E \in \mathscr{F}$$

μ_2在$\mathscr{F}$上定义为

$$\mu_2(\varnothing) = 0, \mu(E) = +\infty, \quad \forall E \neq \varnothing, \ E \in \mathscr{F}$$

这样 μ_1和 μ_2 都是测度。μ_2 既非有限测度也非 σ 有 限 测度。

(2) 设 $(\Omega, \mathscr{F})$ 是可测空间。ω_0 是一固定元素，$\omega_0 \in \Omega$。定义集函数 μ 为

$$\mu(E) = \begin{cases} 0, & \text{若 } \omega_0 \notin E \\ 1, & \text{若 } \omega_0 \in E \end{cases}$$

可以看到μ是有限测度。 这种测度可称为聚集于ω_0的单位测度。

(3) 设 $\Omega = N = \{1, 2, 3, \cdots\}$ 是自然数的集合，$\mathscr{F}$是N的一切子集所成的类，$\mathscr{F}$是σ代数。

$E \in \mathscr{F}$，若E是一个有限集(只含有限个自然数)，定义

$$\mu(E) = E \text{ 中所含自然数的个数}$$

$E \in \mathscr{F}$，若E是一个无限集，定义

$$\mu(E) = +\infty$$

这样定义的μ是测度。它不是有限测度，但它是σ有限测度。

(4) 设 $\Omega = R$，$\mathscr{F} = \mathscr{B}$ (Borel σ代数)，在1.9节中我们将证明$\mathscr{B}$上存在唯一的测度λ，开区间的测度与开区间的长度相等，即若 $E = (a, b)$，那么 $\mu(E) = b - a$。这

种唯一的测度称为 Lebesgue 测度，它不是有限测度，而是 σ 有限测度。

(5) 设 $\Omega = R$，$\mathscr{F} = \mathscr{B}$，$f$ 是单调连续函数. 在 1.9 节中我们将证明 $\mathscr{B}$ 上存在唯一的测度 λ_f，若 $E = (a, b)$ 则

$$\lambda_f(E) = f(b) - f(a)$$

λ_f 称为 Lebesgue-Stieltjes 测度.

引理 1.3.1 设 μ 是定义在 σ 代数 $\mathscr{F}$ 上的测度. 若 $E \in \mathscr{F}$，$F \in \mathscr{F}$ 且 $E \subset F$，那么 $\mu(E) \leqslant \mu(F)$。进一步，若 $\mu(E) < +\infty$，那么 $\mu(F - E) = \mu(F) - \mu(E)$。

证明 因为 $F = E \cup (F - E)$ 且 $E \cap (F - E) = \varnothing$ 因此有

$$\mu(F) = \mu(E) + \mu(F - E) \tag{1.3.2}$$

又因为 $\mu(F - E) \geqslant 0$，于是 $\mu(F) \geqslant \mu(E)$。若 $\mu(E) < +\infty$，在式 (1.3.2) 两边都减去 $\mu(E)$，就得到所需要的结果。 ∎

引理 1.3.1 上半部分所给出的性质，称为测度的单调性。

引理 1.3.2 设 μ 是 σ 代数 $\mathscr{F}$ 上的测度。

(1) 若 $\{E_n\}$ 是 $\mathscr{F}$ 中的增序列，那么

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \lim \mu(E_n) \tag{1.3.3a}$$

称为测度 μ 的上连续性；

(2) 若 $\{F_n\}$ 是 $\mathscr{F}$ 中的减序列且对某一 n，$\mu(F_n) < \infty$，那么

$$\mu\left(\bigcap_{n=1}^{\infty} F_n\right) = \lim \mu(F_n) \tag{1.3.3b}$$

称为测度 μ 的下连续性。性质 (1) 和 (2) 统称为测度的连续性。

证明 若对某一个 n，$\mu(E_n)=+\infty$，那么式（1.3.3a）两边均为 $+\infty$，因此我们可以假定

$$\mu(E_n)<+\infty,\quad \forall n\geqslant 1$$

令 $A_1=E_1$，$A_2=E_2-E_1,\cdots$，$A_n=E_n-E_{n-1},\cdots$ 那么 $\{A_n\}$ 是 $\mathscr{F}$ 中的不相交序列

$$E_n=\bigcup_{j=1}^{n}A_j,\quad \bigcup_{n=1}^{\infty}E_n=\bigcup_{n=1}^{\infty}A_n$$

因为测度 μ 是可列可加的，即

$$\mu\left(\bigcup_{n=1}^{\infty}E_n\right)=\sum_{n=1}^{\infty}\mu(A_n)=\lim_{m\to\infty}\sum_{n=1}^{m}\mu(A_n)$$

由引理 1.3.1 知，$\mu(A_n)=\mu(E_n)-\mu(E_{n-1})$ $\mu(E_0)=\mu(\varnothing)=0$，因此

$$\sum_{n=1}^{m}\mu(A_n)=\mu(E_m)$$

于是式（1.3.3a）成立。

（2）令 $E_n=F_1-F_n$，E_n 是增序列．利用本引理(1)和引理 1.3.1 可得

$$\begin{aligned}\mu\left(\bigcup_{n=1}^{\infty}E_n\right)&=\lim\mu(E_n)\\&=\lim[\mu(F_1)-\mu(F_n)]\\&=\mu(F_1)-\lim\mu(F_n)\end{aligned}$$

由于

$$\bigcup_{n=1}^{\infty}E_n=F_1-\bigcap_{n=1}^{\infty}F_n$$

因此

$$\mu\left(\bigcup_{n=1}^{\infty}E_n\right)=\mu(F_1)-\mu\left(\bigcap_{n=1}^{\infty}F_n\right)$$

组合上述两个方程可得式(1.3.3b)。 ∎

定义 1.3.2 包含Ω和Ω子集所成的σ代数以及定义于σ代数上的测度的三元组$(\Omega, \mathscr{F}, \mu)$称为测度空间。

下面继续介绍与测度有关的术语：

(1) 若存在子集$N \in \mathscr{F}$, $\mu(N)=0$，则命题在$\mathscr{C}(N)$(N的补集)上成立，于是称这个命题几乎处处成立。

(2) 若$N \in \mathscr{F}$，$\mu(N)=0$，f和g是定义在Ω上的两个函数，$f=g$，$\forall \omega \notin N$，则称f和g几乎处处相等，记作

$$f=g, \mu\text{-a. e.}$$

或简单记作

$$f=g, \quad \text{a. e.}$$

(3) 若$\{f_n\}$是Ω上的函数列且

$$\mu\{\lim f_n(\omega) \neq f(\omega)\}=0$$

则称f_n几乎处处收敛于f，并记作

$$f=\lim f_n, \quad \text{a. e.}$$

习　题　1.3

1. 如果μ是$\mathscr{F}$上的测度，A是一给定集，$A \in \mathscr{F}$，λ定义为

$$\lambda(E)=\mu(A \cap E), \ \forall E \in \mathscr{F}$$

试证明λ是$\mathscr{F}$上的测度。这里$\mathscr{F}$是σ代数。

2. 若$\mu_1, \cdots, \mu_n$都是$\mathscr{F}$上的测度，$a_1, \cdots, a_n$均为非负实数，且

$$\lambda(E)=\sum_{j=1}^{n} a_j \mu_j(E), \ \forall E \in \mathscr{F}$$

试证明λ是$\mathscr{F}$上的测度。这里$\mathscr{F}$是σ代数。

3. 若$\{\mu_n\}$是一测度序列，$\mu_n(\Omega)=1$，λ定义为

$$\lambda(E)=\sum_{n=1}^{\infty} 2^{-n} \mu_n(E), \ \forall E \in \mathscr{F}$$

试证明λ是$\mathscr{F}$上的测度且$\lambda(\Omega)=1$($\mathscr{F}$是σ代数)。

4. 设$\Omega=N=\{1,2,3,\cdots\}$，$\mathscr{F}$是由$\Omega$的一切子集所成的类。

若 $\{a_n\}$ 是一非负实数列，μ 定义为

$$\mu(\varnothing)=0$$

$$\mu(E)=\sum_{n\in E}a_n,\quad \forall E\neq\varnothing$$

试证明 μ 是 $\mathscr{F}$ 上的测度（$\mathscr{F}$ 是 σ 代数）。

5. 设 Ω 是一不可数集，$\mathscr{F}$ 是 Ω 一切子集所成的 σ 代数。集函数 μ 定义为

$$\mu(E)=0,\ \text{当 } E \text{ 是可数集}$$

$$\mu(E)=+\infty,\ \text{当 } E \text{ 是不可数集}$$

试证明 μ 是 $\mathscr{F}$ 上的测度。

6. 设 $\Omega=N$，$\mathscr{F}$ 是 N 的一切子集所组成的 σ 代数：

$$\mu(E)=\begin{cases}0, & \text{当 } E \text{ 为有限集}\\ +\infty, & \text{当 } E \text{ 为无限集}\end{cases}$$

μ 是否为测度？

7. 若 $\Omega=N$，$\mathscr{F}$ 是 N 的一切子集所成的 σ 代数，对一切 $E\in\mathscr{F}$，$\lambda(E)=+\infty$。λ 是否为测度？

8. 设 $(\Omega,\mathscr{F},\mu)$ 是测度空间，$\{E_n\}$ 是 $\mathscr{F}$ 中的序列，对某一 n，$\mu(E_n)<\infty$。证明

$$\mu(\liminf E_n)\leqslant\liminf \mu(E_n)$$

[参见式（1.1.5b）]

9. 利用式（1.1.5a）的记号，证明当 $\mu(\bigcup E_n)<+\infty$ 时

$$\limsup \mu(E_n)\leqslant\mu(\limsup E_n)$$

并举例说明当 $\mu(\bigcup E_n)=+\infty$ 时，不等式不成立。

10. 设 $(\Omega,\mathscr{F},\mu)$ 为测度空间，令

$$\mathscr{Z}=\{E\subset\Omega\in\mathscr{F}\mid\mu(E)=0\}$$

$\mathscr{Z}$ 是否为 σ 代数？证明(1)若 $E\in\mathscr{Z}$，$F\in\mathscr{F}$，那么 $E\cap F\in\mathscr{Z}$；(2) 若 $\{E_n\}$ 是 $\mathscr{Z}$ 中的序列，那么 $\bigcup E_n\in\mathscr{Z}$。

11. 设 $(\Omega,\mathscr{F},\mu)$ 为测度空间，$\mathscr{Z}$ 是习题 1.3，10 中定义的集类，并设

$$\mathscr{F}'=\{E\cup Z\mid E\in\mathscr{F},\ z\in\mathscr{Z}\}$$

证明 $\mathscr{F}'$ 是 σ 代数，且称为 $\mathscr{F}$ 的完备化（关于 μ）的 σ 代数。

12. 应用习题 1.3,11 的记号,在 $\mathscr{F}'$ 上定义集函数

$$\mu'(E\cup z)=\mu(E)$$

其中 $E\in\mathscr{F}$, $z\in\mathfrak{Z}$, 试证明 μ' 是 $\mathscr{F}'$ 上的测度,它在 $\mathscr{F}$ 上与 μ 重合,μ' 称为 μ 的完备化,或完备测度。

1.4 积　　分

提要 在本节我们首先引进非负简单可测函数的积分,然后讨论任意非负广义实值可测函数的积分。在本节中,我们还导出了今后常常要用到的单调收敛定理及其重要推论。

可测函数和测度概念的引进为讨论积分准备了条件,在本节我们首先引进非负简单可测函数的积分,然后讨论任意非负广义实值可测函数的积分。为此我们引进如下一些记号:

以下的讨论都是在给定的测度空间 $(\Omega,\mathscr{F},\mu)$ 上进行的。我们用 $M=M(\Omega,\mathscr{F})$ 表示 Ω 到 $\overline{R}$ 的可测函数全体,$M^+=M^+(\Omega,\mathscr{F})$ 表示 Ω 到 $\overline{R}$ 的所有非负可测函数。

我们先讨论 M^+ 类中的函数关于 μ 的积分。简单函数是 Ω 到 R 的函数,它的定义如下:

定义 1.4.1 只取有限个实数值的函数,称为简单函数。

简单可测函数可表示为下列形式:

$$\varphi=\sum_{j=1}^{n}a_j\chi_{E_j}\tag{1.4.1}$$

其中 $a_j\in R$, χ_{E_j} 表示 E_j 的示性函数,$E_j\in\mathscr{F}$。在 φ 的式 (1.4.1) 表达式中存在唯一的标准表达式。在标准表达式中 E_j 是不相交集合,a_j 是不同的实数。显然,如果 $a_1,\cdots,a_n$

是 n 个不同的值，令 $E_j=\{\omega\in\Omega\,|\,\varphi(\omega)=a_j\}$，那么 E_j 不相交且 $\bigcup_{j=1}^{n} E_j=\Omega$。

当然，如果我们不要求 a_j 不相同，E_j 不相交，那么简单函数作为示性函数的线性组合，有许多表示式。

定义 1.4.2 若 φ 是简单函数，$\varphi\in M^+(\Omega,\mathscr{F})$，则式 (1.4.1) 是 φ 的标准表达式，于是 φ 关于 μ 的积分定义为一个广义实数

$$\sum_{j=1}^{n} a_j\mu(E_j)$$

记作

$$\int\varphi d\mu=\sum_{j=1}^{n} a_j\mu(E_j) \tag{1.4.2}$$

在表达式 (1.4.2) 中我们约定当零乘以 $+\infty$ 时等于零，即 $0(+\infty)=0$。同时我们知道 $\varphi\in M^+$ 时 a_j 是非负实数，因此不可能出现 $(+\infty)-(+\infty)$ 这种不确定的运算。因此这样的定义是合理的，完全确定的。

下面讨论积分的基本性质：

引理 1.4.1 (1) 若 φ 和 ϕ 是两个简单函数，且 $\varphi,\phi\in M^+(\Omega,\mathscr{F})$，$c\geqslant 0$，那么

$$\int c\varphi d\mu=c\int\varphi d\mu$$

$$\int(\varphi+\phi)d\mu=\int\varphi d\mu+\int\phi d\mu \tag{1.4.3}$$

(2) λ 是 $\mathscr{F}$ 上的集函数，并定义为

$$\lambda(E)=\int\varphi\chi_E d\mu,\quad\forall E\in\mathscr{F}$$

那么 λ 是 $\mathscr{F}$ 上的测度。

证明 若 $c=0$，那么 $c\varphi\equiv 0$，引理中的等式成立。

若 $c>0$，那么 $c\varphi\in M^+$，其标准表达式为

$$c\varphi=\sum_{j=1}^{n}ca_j\chi_{E_j}$$

[因为 φ 的标准表达式为式(1.4.1)].

$$\int c\varphi d\mu=\sum_{j=1}^{n}ca_j\mu(E_j)$$

$$=c\sum_{j=1}^{n}a_j\mu(E_j)=c\int\varphi d\mu$$

设 φ 和 ψ 的标准表达式分别为

$$\varphi=\sum_{j=1}^{n}a_j\chi_{E_j},\quad \psi=\sum_{k=1}^{m}b_k\chi_{F_k}$$

然后把 $\varphi+\psi$ 作为不相交集 $E_j\cap F_k$ 的示性函数的线性组合，它不一定是标准表达式，因为此时 a_j+b_k 不一定不同．设 c_h，$h=1,\cdots,p$，表示 $\{a_j+b_k:\ j=1,\cdots,n,\ k=1,\cdots,m\}$ 中 p 个不同的值：

$$G_h=\bigcup_{a_j+b_k=c_h}(E_j\cap F_k)$$

这样

$$\mu(G_h)=\sum_{(h)}\mu(E_j\cap F_k)$$

这里 (h) 表示使得 $a_j+b_k=c_h$ 的那些 j 与 k。因此 $\varphi+\psi$ 的标准表达式为

$$\varphi+\psi=\sum_{h=1}^{p}c_h\chi_{G_h}$$

根据定义 1.4.1

$$\int(\varphi+\psi)d\mu=\sum_{h=1}^{p}c_h\mu(G_h)=\sum_{h=1}^{p}\sum_{(h)}c_h\mu(E_j\cap F_k)$$

$$= \sum_{h=1}^{p} \sum_{(h)} (a_j + b_k)\mu(E_j \cap F_k)$$

$$= \sum_{j=1}^{n} \sum_{k=1}^{m} (a_j + b_k)\mu(E_j \cap F_k)$$

$$= \sum_{j=1}^{n} \sum_{k=1}^{m} a_j \mu(E_j \cap F_k)$$

$$+ \sum_{j=1}^{n} \sum_{k=1}^{m} b_k \mu(E_j \cap F_k)$$

因为

$$\Omega = \bigcup_{j=1}^{n} E_j = \bigcup_{k=1}^{m} F_k$$

所以

$$\mu(E_j) = \sum_{k=1}^{m} \mu(E_j \cap F_k)$$

$$\mu(F_k) = \sum_{j=1}^{n} \mu(E_j \cap F_k)$$

利用求和号交换得

$$\int (\varphi + \psi) d\mu$$

$$= \sum_{j=1}^{n} a_j \mu(E_j) + \sum_{k=1}^{m} b_k \mu(F_k)$$

$$= \int \varphi d\mu + \int \psi d\mu$$

为了证明(2)，考虑

$$\varphi \chi_E = \sum_{j=1}^{n} a_j \chi E_j \cap E$$

利用我们已经证明的结论，可得

$$\lambda(E)=\int\varphi\chi_E d\mu=\sum_{j=1}^{n}a_j\int\chi_{E_j\cap E}d\mu$$

$$=\sum_{j=1}^{n}a_j\mu(E_j\cap E)$$

根据习题 1.3 的 1，2 可知 $\lambda(E)$ 是$\mathscr{F}$上的测度。 ∎

设 Ω 为某一有限区间，于是简单函数 $\varphi(\varphi\in M^{+})$ 的积分可以用有限个"矩形"的面积和来表示. （抽象积分不一定是面积，但用面积来作解释可以帮助理解，如图 1.4.1 所示）

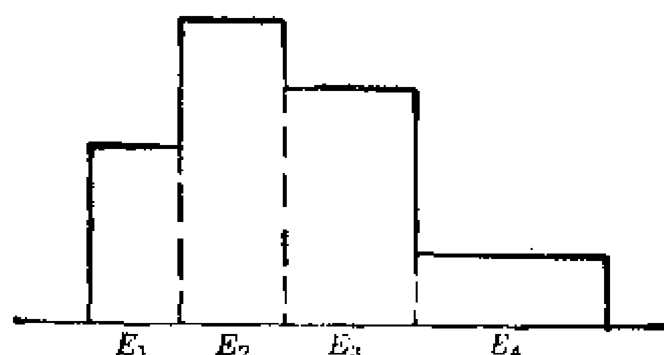


图 1.4.1 非负可测简单函数的积分示意图

这样，非负可测函数 f 的积分，就可以用小于或等于 f 的 φ，$\varphi\in M^{+}$ 的积分的上确界来定义，如图 1.4.2 所示。

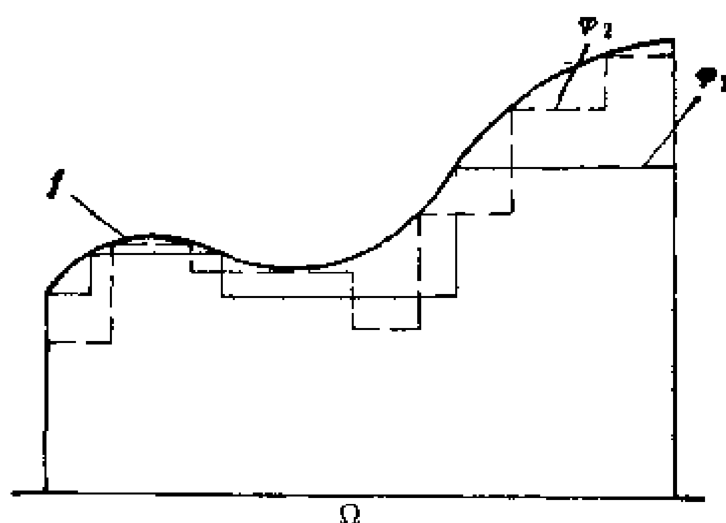


图 1.4.2 非负可测函数积分的示意图

现在我们引进 M^+ 中任意函数的积分，把它定义为 φ 的积分的上确界，$\varphi\leqslant f$，注意这里并不要求积分值有限。

定义 1.4.3 若 $f\in M^+(\Omega, \mathscr{F})$，$f$ 关于 μ 的积分定义为

$$\int f d\mu = \sup_{\varphi\leqslant f}\int \varphi d\mu$$

这里 $\varphi\in M^+$ 表示简单函数，上确界在满足 $0\leqslant\varphi(\omega)\leqslant f(\omega)$ 的范围内取值。若 $f\in M^+$，$E\in\mathscr{F}$，那么 $f\chi_E\in M^+$。f 关于 μ 在 E 上的积分定义为

$$\int_E f d\mu = \int f\chi_E d\mu \tag{1.4.4}$$

现在证明积分的单调性。

引理 1.4.2 (1) 若 f 和 $g\in M^+(\Omega, \mathscr{F})$ 且 $f\leqslant g$，那么

$$\int f d\mu \leqslant \int g d\mu \tag{1.4.5}$$

(2) 若 $f\in M^+(\Omega, \mathscr{F})$，$E, F\in\mathscr{F}$ 且 $E\subset F$，那么

$$\int_E f d\mu \leqslant \int_F f d\mu$$

证明 (1) 设 $\varphi\in M^+$，φ 是简单函数，那么当 $f\leqslant g$ 时

$$\sup_{0\leqslant\varphi\leqslant f}\int \varphi d\mu \leqslant \sup_{0\leqslant\varphi\leqslant g}\int \varphi d\mu$$

因此式(1.4.5)成立。

(2) 因为 $f\chi_E\leqslant f\chi_F$，利用本引理(1)得证。 ▌

下面介绍 B. Levy 引进的重要结果——单调收敛定理。这个定理是研究抽象 Lebesgue 积分收敛性的关键。

定理 1.4.1 (单调收敛定理) 若 $\{f_n\}$ 是 $M^+(\Omega, \mathscr{F})$中的单调增序列，$f_n\to f$ (当 $n\to\infty$)，那么

$$\int f d\mu = \lim \int f_n d\mu \tag{1.4.6}$$

证明 首先根据推论 1.2.2，可知 f 可测。因为 $f_n \leqslant f_{n+1} \leqslant f$，由引理 1.4.2 (1)得

$$\int f_n d\mu \leqslant \int f_{n+1} d\mu \leqslant \int f d\mu, \forall n = 1,2,\cdots$$

所以有

$$\lim \int f_n d\mu \leqslant \int f d\mu$$

为了建立反向不等式，设 α 是满足下式的实数，$0<\alpha<1$，并设 φ 是简单可测函数，满足 $0 \leqslant \varphi \leqslant f$。令

$$A_n = \{\omega \in \Omega: f_n(\omega) \geqslant \alpha\varphi(\omega)\}$$

因此

$$A_n \in \mathscr{F}, \ A_n \subset A_{n+1}, \ 且 \ \Omega = \cup A_n$$

由引理 1.4.2 可知

$$\int_{A_n} \alpha\varphi d\mu \leqslant \int_{A_n} f_n d\mu \leqslant \int f_n d\mu \tag{1.4.7}$$

因为 $\{A_n\}$ 是单调增序列，$\cup A_n = \Omega$。根据引理 1.4.1 (2)和引理 1.3.2 (1)得

$$\int \varphi d\mu = \lim \int_{A_n} \varphi d\mu$$

所以对于式(1.4.7)两边取极限得

$$\alpha \int \varphi d\mu \leqslant \lim \int f_n d\mu$$

因为上式对于 $0<\alpha<1$ 均成立，因此

$$\int \varphi d\mu \leqslant \lim \int f_n d\mu$$

又因为 φ 是满足 $0 \leqslant \varphi \leqslant f$ 的任意可测简单函数，因此

$$\int f d\mu = \sup_{\varphi} \int \varphi d\mu \leqslant \lim \int f_n d\mu$$

把两个不等式组合起来，我们便得到了式(1.4.6)。 ▮

注意： 这里我们并没有假定式(1.4.6)两边有限． 显然 $\int f_n d\mu$ 是单调增序列，它在 $\bar{R}$ 中总是有极限的，但在 R 中就不一定了。

下面给出单调收敛定理的一些推论：

推论 1.4.1 (1) 若 $f\in M^+$，$c\geqslant 0$，那么 $cf\in M^+$ 且

$$\int cf d\mu = c\int f d\mu$$

(2) 若 $f, g\in M^+$，那么 $f+g\in M^+$，且

$$\int (f+g)d\mu = \int f d\mu + \int g d\mu$$

证明 (1) 若 $c=0$，结论显然成立．若 $c>0$，$\varphi_n\in M^+$ 是单调增简单函数列，在 Ω 上收敛于 f (参见引理 1.2.5)．那么 $\{c\varphi_n\}$ 也是一单调增简单函数列，收敛于 cf。 应用引理 1.4.1(1) 和单调收敛定理得

$$\begin{aligned}\int cf d\mu &= \lim\int c\varphi_n d\mu \\ &= c\lim\int \varphi_n d\mu \\ &= c\int f d\mu\end{aligned}$$

(2) 若 $\{\varphi_n\}$，$\{\phi_n\}$ 是两个单调增简单函数列，分别收敛于 f 和 g，那么 $\{\varphi_n+\phi_n\}$ 也是一单调增序列并收敛于 $f+g$。 利用引理 1.4.1(1) 和单调收敛定理可得

$$\begin{aligned}\int (f+g)d\mu &= \lim\int (\varphi_n+\phi_n)d\mu \\ &= \lim\int \varphi_n d\mu + \lim\int \phi_n d\mu \\ &= \int f d\mu + \int g d\mu\end{aligned}$$

▮

下面是单调收敛定理的推论（Fatou 引理），它在应用中与单调收敛定理同样重要。

引理 1.4.3 (Faton 引理) 若$\{f_n\}$是 M^+ 中的序列，那么

$$\int(\liminf f_n)d\mu \leqslant \liminf \int f_n d\mu \tag{1.4.8}$$

证明 设 $g_m=\inf\{f_m, f_{m+1}, \cdots\}$，所以 $g_m \leqslant f_n$，$\forall m \leqslant n$。由积分的单调性，有

$$\int g_m d\mu \leqslant \int f_n d\mu, \quad \forall m \leqslant n$$

因此

$$\int g_m d\mu \leqslant \liminf \int f_n d\mu$$

因为 $\{g_m\}$ 是单调增序列，且 $g_m \to \liminf f_n$。

由单调收敛定理可知

$$\int(\liminf f_n)d\mu = \lim \int g_m d\mu$$

$$\leqslant \liminf \int f_n d\mu$$ ∎

在习题中将看到，若 $f_n \geqslant 0$ 不成立，Fatou 引理也可能不成立。

推论 1.4.2 若 $f \in M^+$，λ 是定义在 $\mathscr{F}$ 上的集函数：

$$\lambda(E) = \int_E f d\mu \tag{1.4.9}$$

那么 λ 是测度。

证明 因为 $f \geqslant 0$，因此 $\lambda(E) \geqslant 0$。若 $E=\varnothing$，$f\chi_E=0$，则 $\lambda(\varnothing)=0$。为了证明 λ 的可列可加性，设 $\{E_n\}$ 是 $\mathscr{F}$ 中不相交的序列，且 $\bigcup_{n=1}^{\infty} E_n = E$。令

$$f_n = \sum_{k=1}^{n} f\chi_{E_k}$$

由推论 1.4.1(2) 可知

$$\int f_n d\mu = \sum_{k=1}^{n} \int f\chi_{E_k} d\mu = \sum_{k=1}^{n} \lambda(E_k)$$

因为 f_n 是 M^+ 中的单调增序列，$f_n \to f\chi_E$，由单调收敛定理可得

$$\lambda(E) = \int f\chi_E d\mu = \lim \int f_n d\mu = \sum_{k=1}^{\infty} \lambda(E_k)$$ ∎

推论 1.4.3 假设 $f \in M^+$，当且仅当 $f = 0$，a. e. 时

$$\int f d\mu = 0 \tag{1.4.10}$$

证明 若式(1.4.10)成立，令

$$E_n = \left\{\omega \in \Omega: \ f(\omega) > \frac{1}{n}\right\}$$

于是

$$f \geqslant \frac{1}{n} \chi_{E_n}$$

由此得

$$0 = \int f d\mu \geqslant \frac{1}{n} \mu(E_n) \geqslant 0$$

于是有 $\mu(E_n) = 0$；因为

$$\{\omega \in \Omega: \ f(\omega) > 0\} = \bigcup_{n=1}^{\infty} E_n$$

$$\mu\{\omega \in \Omega: \ f(\omega) > 0\} \leqslant \sum_{n=1}^{\infty} \mu(E_n) = 0$$

因此

$$f(\omega) = 0, \text{ a. e.}$$

反之，若 $f(\omega)=0$, a. e., 即令

$$E=\{\omega\in\Omega:\ f(\omega)>0\}$$

则 $\mu(E)=0$。令 $f_n=n\chi_E$，可得 $f\leqslant\liminf f_n$。由 Fatou 引理，有

$$0\leqslant\int f d\mu\leqslant\liminf\int f_n d\mu=0$$ ▮

推论 1.4.4 假定 $f\in M^+$，λ 是由式(1.4.9)定义的集函数，那么若 $E\in\mathscr{F}$，$\mu(E)=0$ 可推出 $\lambda(E)=0$。我们称 λ 关于 μ 绝对连续，并记作 $\lambda\ll\mu$。

证明 若 $\mu(E)=0$，$\forall E\in\mathscr{F}$，那么 $f\chi_E=0$, a. e. 由推论 1.4.3 可得

$$\lambda(E)=\int f\chi_E d\mu=0$$ ▮

现在我们证明把收敛换成几乎处处收敛，单调收敛定理仍然成立。

推论 1.4.5 $\{f_n\}$ 是 M^+ 中的单调增序列，$\lim f_n=f$ a. e.，那么

$$\int f d\mu=\lim\int f_n d\mu$$

证明 设 $N\in\mathscr{F}$ 且 $\mu(N)=0$，f_n 在 $M=\Omega-N$ 上点点收敛于 f。那么 $f_n\chi_M\to f\chi_M$。根据单调收敛定理

$$\int f\chi_M d\mu=\lim\int f_n\chi_M d\mu$$

因为 $\mu(N)=0$，因此 $f\chi_N=0$, a.e., $f_n\chi_N=0$, a.e.，由推论 1.4.3 得

$$\int f\chi_N d\mu=0,\quad\int f_n\chi_N d\mu=0$$

由于 $f=f\chi_M+f\chi_N$，$f_n=f_n\chi_M+f_n\chi_N$，因此有

$$\int f d\mu=\int f\chi_M d\mu=\lim\int f_n\chi_M d\mu=\lim\int f_n d\mu$$ ▮

推论 1.4.6 设 $\{g_n\}$ 是 M^+ 中一序列，那么

$$\int\left(\sum_{n=1}^{\infty} g_n\right) d\mu = \sum_{n=1}^{\infty}\left(\int g_n d\mu\right)$$

证明 令 $f_n = g_1 + \cdots + g_n$，f_n 是单调增序列，利用单调收敛定理，推论得证。 ∎

习 题 1.4

1. 设 $\varphi \in M^+(\Omega, \mathscr{F})$ φ 是简单函数，它具有下列表达式（并非标准表达式）：

$$\varphi = \sum_{k=1}^{m} b_k \chi_{E_k}$$

其中 $b_k \in R$，$E_k \in \mathscr{F}$. 证明

$$\int \varphi d\mu = \sum_{k=1}^{m} b_k \mu(E_k)$$

2. 证明简单函数的和，简单函数与数相乘，简单函数之积都是简单函数. 换句话说，$M(\Omega, \mathscr{F})$ 中的简单函数组成 $M(\Omega, \mathscr{F})$ 的一个向量子空间.

3. 若 φ_1 和 φ_2 都是 $M(\Omega, \mathscr{F})$ 中的简单函数，证明

$$\psi = \sup\{\varphi_1, \varphi_2\}, \ \eta = \inf\{\varphi_1, \varphi_2\}$$

也都是 $M(\Omega, \mathscr{F})$ 中的简单函数.

4. 若 $f \in M^+$，且 $C>0$，那么 $\varphi \to \psi = C\varphi$ 是 $\{\varphi \in M^+: \varphi \leqslant f\}$ 到 $\{\psi \in M^+: \psi \leqslant cf\}$ 的一一对应. 利用这一关系证明推论 1.4.1(1).

5. 令 $\Omega = N$，$\mathscr{F}$ 是N一切子集所成的 σ 代数 $E \in \mathscr{F}$，$\mu(E) = E$ 中自然数个数，设 f 是N上的非负函数，证明 $f \in M^+(\Omega, \mathscr{F})$，且

$$\int f d\mu = \sum_{n=1}^{\infty} f(n)$$

6. 设 $\Omega = R$，$\mathscr{F} = \mathscr{B}$，$\lambda$ 是 $\mathscr{B}$ 上的 Lebesgue 测度. 令 $f_n = \chi_{[0,n]}$，则 f_n 是一个单调增序列，$f_n \to f = \chi_{[0,+\infty)}$. 显然这个函数列是一致有界的，界为 1. f_n 的所有积分均有限，但

$$\int f d\lambda = +\infty$$

试问能否应用单调收敛定理？

7. 设 $\Omega=R$，$\mathscr{F}=\mathscr{B}$，λ 是 $\mathscr{B}$ 上的 Lebesgue 测度．令 $f_n=\frac{1}{n}\chi_{[n,+\infty)}$，证明 $\{f_n\}$ 是单调减序列，$f_n\to f=0$，但是

$$0=\int f d\lambda \neq \lim\int f_n d\lambda=+\infty$$

因此，对于 M^+ 的递减序列并没有相应的单调收敛定理。

8. 令 $f_n=\frac{1}{n}\chi_{[0,n]}$，$f=0$．证明 (i) f_n 均匀收敛于 f．(ii) $\int f d\lambda \neq \lim\int f_n d\lambda$．为什么这个事实与单调收敛定理不矛盾？能否在本题中应用 Fatou 引理？

(2) 令 $g_n=n\chi_{[\frac{1}{n},\frac{2}{n}]}$，$g=0$．证明

$$\int g d\lambda \neq \lim\int g_n d\lambda$$

g_n 是否均匀收敛到 g？能否在本题中应用单调收敛定理？能否应用 Fatou 引理？

9. 设 $(\Omega,\mathscr{F},\mu)$ 是有限测度空间，$\{f_n\}$ 是 $M^+(\Omega,\mathscr{F})$ 中的可测实函数，f_n 均匀收敛于 f，证明

(1) $f\in M^+(\Omega,\mathscr{F})$；

(2) $\int f d\mu=\lim\int f_n d\mu$．

10. 设 $\Omega=[a,b]$ 是一有限闭区间，$\mathscr{B}$ 是 Ω 中 Borel 集所成的类，λ 是 $\mathscr{B}$ 上 Lebesgue 测度．若 f 是 Ω 上非负连续函数，证明

$$\int f d\lambda=\int_a^b f(x)dx$$

其中等式右端是 f 的 Riemann 积分．（提示：先对非负阶梯函数证明有上式成立，区间示性函数的线性组合称为阶梯函数）

11. 设 $\Omega=[0,+\infty)$，$\mathscr{B}$ 是 Ω 的 Borel 子集所成的类，并由习题 1.3，11 完备化．λ 是 $\mathscr{B}$ 上 Lebesgue 测度．若 f 是 Ω 上非负连续函数，证明

$$\int f d\lambda = \lim_{b\to\infty}\int_a^b f(x)dx$$

若 f 是非负连续函数，证明 Lebesgue 积分与 Riemann 积分重合。

12. 若 $f\in M^+(\Omega, \mathscr{F})$，且

$$\int f d\mu < +\infty$$

证明 $\mu\{\omega\in\Omega: f(\omega) = +\infty\} = 0$。（提示：令 $E_n = \{\omega\in\Omega: f(\omega)\geqslant n\}$，那么 $n\chi_{E_n}\leqslant f$）

13. 若 $f\in M^+(\Omega, \mathscr{F})$，且

$$\int f d\mu < +\infty$$

令 $N = \{\omega\in\Omega: f(\omega)>0\}$，证明 N 是 σ 有限集（即存在一个 $\mathscr{F}$ 中的序列 $\{F_n\}$，$N\subset\bigcup F_n$，且 $\mu(F_n)<+\infty$）。

14. 若 $f\in M^+(\Omega, \mathscr{F})$，且

$$\int f d\mu < +\infty$$

证明对任何 $\varepsilon>0$，存在 $E\in\mathscr{F}$ 使得 $\mu(E)<+\infty$，且

$$\int f d\mu \leqslant \int_E f d\mu + \varepsilon$$

15. 假设 $\{f_n\}\subset M^+(\Omega, \mathscr{F})$，$f_n\to f$，且

$$\int f d\mu = \lim\int f_n d\mu < +\infty$$

证明

$$\int_E f d\mu = \lim\int_E f_n d\mu, \quad \forall E\in\mathscr{F}$$

16. 证明当条件

$$\lim\int f_n d\mu < +\infty$$

被破坏时，习题 1.4，15 的结论可能不成立。

1.5 可 积 函 数

提要 本节讨论一般可测函数的积分，并证明

积分论中的又一著名定理——控制收敛定理。本节后半部分讨论积分号下取极限的条件。

在 1.4 节中，我们定义了非负可测函数的积分，这种积分可以等于$+\infty$。在本节我们将讨论能取正值和负值的一般可测函数的积分，而且要求函数的积分等于一个有限数。

定义 1.5.1 可测函数 f 的正部 f^+，负部 f^- 的积分(关于 μ) 均有限，则称 f 可积，f 关于 μ 的积分定义为

$$\int f d\mu = \int f^+ d\mu - \int f^- d\mu \tag{1.5.1}$$

若 $E \in \mathscr{F}$，定义

$$\int_E f d\mu = \int_E f^+ d\mu - \int_E f^- d\mu$$

Ω 上 $\mathscr{F}$ 可测的可积函数的全体组成 $L = L(\Omega, \mathscr{F}, \mu)$。

虽然 f 的积分定义为正部 f^+ 积分与负部 f^- 积分之差，但容易看出：如果 $f = f_1 - f_2$，f_1 和 f_2 都是非负可测函数，且都具有有限积分，那么

$$\int f d\mu = \int f_1 d\mu - \int f_2 d\mu$$

事实上，因为 $f^+ - f^- = f_1 - f_2$，于是 $f^+ + f_2 = f^- + f_1$。利用推论 1.4.1(2)，可得

$$\int f^+ d\mu + \int f_2 d\mu = \int f_1 d\mu + \int f^- d\mu$$

由于上述四项都是有限的，因此

$$\int f d\mu = \int f^+ d\mu - \int f^- d\mu = \int f_1 d\mu - \int f_2 d\mu.$$

设 $(\Omega, \mathscr{F})$ 是可测空间，λ 是定义在 $\mathscr{F}$ 上的广义实值集函数，如果 λ 满足如下条件：

(1) $\lambda(\varnothing) = 0$;

(2) 除有限值外，在 $\pm\infty$ 二值中至多只能取一个值；

(3) 具有可列可加性：当 $\{E_n\}$ 是 $\mathscr{F}$ 中的不相交的集时，有

$$\lambda\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \lambda(E_n)$$

则称 λ 是 $(\Omega, \mathscr{F})$ 上一个广义测度。

例 设 p_1 与 p_2 是 $(\Omega, \mathscr{F})$ 上的两个测度，其中至少有一个是有限测度，于是

$$\mu(A) = p_1(A) - p_2(A), \quad \forall A \in \mathscr{F}$$

是 $(\Omega, \mathscr{F})$ 上的广义测度。

引理 1.5.1 若 $f \in L$，λ 是定义在 $\mathscr{F}$ 上的实值集函数

$$\lambda(E) = \int_E f d\mu \tag{1.5.2}$$

则 λ 是 $\mathscr{F}$ 上的广义测度。

证明 因为 f^+ 和 $f^- \in M^+$，根据推论 1.4.1 由下式定义的集函数：

$$\lambda^+(E) = \int_E f^+ d\mu, \lambda^-(E) = \int_E f^- d\mu$$

是 $\mathscr{F}$ 上的测度。又因为 $f \in L$，$\lambda^+(E)$ 和 $\lambda^-(E)$ 均有限。故 $\lambda = \lambda^+ - \lambda^-$ 是广义测度。 ▮

由式(1.5.2)定义的集函数也称为 f 关于 μ 的不定积分。

由于 λ 是广义测度，若 $\{E_n\}$ 是 $\mathscr{F}$ 中的不相交序列，且

$$\bigcup_{n=1}^{\infty} E_n = E,$$

那么

$$\int_E f d\mu = \sum_{n=1}^{\infty} \int_{E_n} f d\mu$$

这一式子表明，L 中函数的不定积分是可列可加的。

下面的结果有时称为 Lebesgue 积分的绝对可积性。

定理 1.5.1 可测函数 f Lebesgue 可积的充分必要条件是 $|f|$ Lebesgue 可积。此时

$$\left|\int f d\mu\right| \leqslant \int |f| d\mu \tag{1.5.3}$$

证明 由 f Lebesgue 可积的定义知：$f \in L \Longleftrightarrow f^+$，$f^- \in M^+$ 且均有有限积分。
因为

$$|f| = f^+ + f^- = |f|^+$$

$$|f|^- = 0$$

利用引理 1.4.2(1) 和推论 1.4.1 (2)得

$$\left|\int f d\mu\right| = \left|\int f^+ d\mu - \int f^- d\mu\right|$$

$$\leqslant \int f^+ d\mu + \int f^- d\mu = \int |f| d\mu$$ ∎

推论 1.5.1 若 f 是可测函数，g 是可积函数且

$$|f| \leqslant |g|$$

那么 f 必定可积，且

$$\int |f| d\mu \leqslant \int |g| d\mu$$

证明 由引理 1.4.2 (1)及定理 1.5.1 可证。 ∎

现在我们证明积分在 L 上是线性的，即

定理 1.5.2 若 $f, g \in L$，α 为某常数，那么 $\alpha f \in L$，$f + g \in L$，且有

$$\int \alpha f d\mu = \alpha \int f d\mu$$

$$\int (f + g) d\mu = \int f d\mu + \int g d\mu$$

证明 若 $\alpha = 0$，则 $\alpha f = 0$，因此

$$\int \alpha f d\mu = 0 = \alpha \int f d\mu$$

若 $\alpha > 0$，那么 $(\alpha f)^+ = \alpha f^+$，$(\alpha f)^- = \alpha f^-$，因此

$$\int \alpha f d\mu = \int \alpha f^+ d\mu - \int \alpha f^- d\mu$$

$$= \alpha \left\{\int f^+ d\mu - \int f^- d\mu\right\} = \alpha \int f d\mu$$

对于 $\alpha<0$，可作类似证明。

若 f，$g \in L$，则 $|f|, |g| \in L$。因为 $|f+g| \leqslant |f| + |g|$，由推论 1.4.1 和 1.5.1 可知 $f+g \in L$。为了建立我们所要证明的关系式，考虑

$$f + g = (f^+ + g^+) - (f^- + g^-)$$

因为 $f^+ + g^+$ 和 $f^- + g^-$ 都是非负可积函数，由定义 1.5.1 得

$$\int (f+g) d\mu = \int (f^+ + g^+) d\mu - \int (f^- + g^-) d\mu$$

利用推论 1.4.1 (2)，并重新排列这些项得到

$$\int (f+g) d\mu = \int f^+ d\mu - \int f^- d\mu + \int g^+ d\mu - \int g^- d\mu = \int f d\mu + \int g d\mu$$ ∎

下面给出关于可积函数最重要的收敛定理。

定理 1.5.3 (Lebesgue 控制收敛定理) 设 $\{f_n\}$ 是一可积函数列，它几乎处处收敛于实值可测函数 f。如果存在可积函数 g，使得 $|f_n| \leqslant g$，$\forall n$。那么 f 必定可积且

$$\int f d\mu = \lim \int f_n d\mu \tag{1.5.4}$$

证明 在零测集 N 上重新定义 f_n 和 f，在 $\Omega - N$ 上，f_n，f 按原定义，那么

$$f_n \to f, \quad \forall \omega \in \Omega$$

由推论 1.5.1 可知 f 可积。由于 $g + f_n \geqslant 0$，由 Fatou 引理和定理 1.5.2 可得

$$\begin{aligned}
\int g d\mu + \int f d\mu = \int (f + g) d\mu \\
&\leqslant \liminf \int (g + f_n) d\mu \\
&= \liminf \left(\int g d\mu + \int f_n d\mu \right) \\
&= \int g d\mu + \liminf \int f_n d\mu
\end{aligned}$$

因为 $\int g d\mu$ 有限，所以

$$\int f d\mu \leqslant \liminf \int f_n d\mu \tag{1.5.5}$$

又因为 $g - f_n \geqslant 0$，再利用 Fatou 引理可得

$$\begin{aligned}
\int g d\mu - \int f d\mu = \int (g - f) d\mu &\leqslant \liminf \int (g - f_n) d\mu \\
&= \int g d\mu - \limsup \int f_n d\mu
\end{aligned}$$

由 $\int g d\mu$ 有限，可得

$$\limsup \int f_n d\mu \leqslant \int f d\mu \tag{1.5.6}$$

于是由式(1.5.5)和(1.5.6)得

$$\int f d\mu = \lim \int f_n d\mu \qquad \blacksquare$$

从这里开始，如果不加说明，我们都假定 f 是 $\Omega \times [a, b]$ 到 R 的函数。对于每一个 $t \in [a, b]$，$f(\omega, t)$ 是可测函数。

在许多应用问题中，需要考虑被积函数含参数的情况．作为 Lebesgue 控制收敛定理的推论我们有：

推论 1.5.2 若对于 $t_0 \in [a, b]$，有

$$f(\omega, t_0) = \lim_{t \to t_0} f(\omega, t), \ \forall \omega \in \Omega \qquad (1.5.7)$$

且存在可积函数 g 使得

$$|f(\omega, t)| \leqslant g$$

那么

$$\int f(\omega, t_0) d\mu = \lim_{t \to t_0} \int f(\omega, t) d\mu$$

证明 设 $\{t_n\}$ 是 $[a,b]$ 中的一个实数序列，$t_n \to t_0$，定义 $f_n(\omega) = f(\omega, t_n)$，$\forall \omega \in \Omega$. 在 $\{f_n\}$上使用 Lebesgue 控制收敛定理，可得所需结论。 ▮

推论 1.5.3 若 $f(\omega, t)$ 关于 t 是连续函数，在 Ω 上存在可积函数 g 使得 $|f(\omega,t)| \leqslant g$. $F(t)$ 定义为

$$F(t) = \int f(\omega, t) d\mu \qquad (1.5.8)$$

那么 $F(t)$ 是 $[a,b]$ 上的连续函数。

证明 利用推论 1.5.2 可以直接得出。 ▮

推论 1.5.4 假定对某一 $t_0 \in [a, b]$ 有 $f(\omega, t_0)$ 关于 μ 可积. 还假定 $\partial f/\partial t$ 在 $\Omega \times [a,b]$ 上存在，且存在一可积函数 g，使得

$$\left|\frac{\partial f}{\partial t}(\omega, t)\right| \leqslant g(\omega)$$

那么由式(1.5.8)定义的 $F(t)$ 在 $[a,b]$ 上可微，且

$$\frac{dF(t)}{dt} = \frac{d}{dt}\int f(\omega,t) d\mu = \int \frac{\partial f}{\partial t}(\omega, t) d\mu$$

证明 设 $t \in [a,b]$ 的任意点，$\{t_n\} \subset [a,b]$，$t_n \neq t$ 且 $t_n \to t$，那么

$$\frac{\partial f}{\partial t}(\omega, t) = \lim \frac{f(\omega, t_n) - f(\omega, t)}{t_n - t}, \ \forall \omega \in \Omega$$

由此可知 $\partial f/\partial t$ 是 Ω 上的可测函数。

若对 $\omega \in \Omega, t \in [a,b]$ 应用中值定理，存在 $t_0 < t_1 < t$

使得

$$f(\omega,t)-f(\omega,t_0)=(t-t_0)\frac{\partial f}{\partial t}(\omega,s_1)$$

由此得

$$|f(\omega,t)|\leqslant|f(\omega,t_0)|+|t-t_0|g(\omega)$$

因为 $f(\omega,t_0),g(\omega)$ 均可积，因此 $f(\omega,t)$ 可积。

若 $t_n\to t$，则

$$\frac{F(t_n)-F(t)}{t_n-t}=\int\frac{f(\omega,t_n)-f(\omega,t)}{t_n-t}\,d\mu$$

中的被积函数被可积函数 g 控制，即

$$\left|\frac{f(\omega,t_n)-f(\omega,t)}{t_n-t}\right|\leqslant g(\omega)$$

利用控制收敛定理，就可得到所需的结论。

推论 1.5.5 在推论 1.5.3 的假设下

$$\int_a^b F(t)dt=\int_a^b\left[\int f(\omega,t)d\mu\right]dt$$
$$=\int\left[\int_a^b f(\omega,t)dt\right]d\mu$$

这里关于 t 的积分 $\int_a^b$ 是 Riemann 意义下的积分。

证明 对于 Riemann 积分，若 φ 是 $[a,b]$ 上的连续函数，那么

$$\frac{d}{dt}\int_a^t\varphi(s)ds=\varphi(t),\quad\forall a\leqslant t\leqslant b$$

设 h 是定义在 $\Omega\times[a,b]$ 上的函数：

$$h(\omega,t)=\int_a^t f(\omega,s)ds$$

因为 Riemann 积分是 Riemann 和的极限，因此对每一个 $t\in[a,b]$，$h(\omega,t)$ 是 ω 的可测函数。

更进一步，因为 $|f(\omega,t)|\leqslant g(\omega)$，因此有

$$|h(\omega,t)| = \left|\int_0^t f(\omega, s)ds\right|$$
$$\leqslant \int_a^t |f(\omega,s)|ds$$
$$\leqslant g(\omega)(b-a)$$

因此，对于每一个 $t\in[a,b]$，$h(\omega,t)$ 是关于 μ 的可积函数.

设 $H(t)$ 是定义在 $[a,b]$ 上的函数，

$$H(t) = \int h(\omega,t)d\mu$$

由推论 1.5.4 可知

$$\frac{dH(t)}{dt} = \int \frac{\partial h(\omega,t)}{\partial t}\, d\mu = \int f(\omega,t)d\mu = F(t)$$

所以

$$\int_a^b F(t)dt = H(b) - H(a)$$
$$= \int [h(\omega,b) - h(\omega,a)]d\mu$$
$$= \int\left[\int_a^b f(\omega,t)dt\right] d\mu$$

即

$$\int_a^b\left[\int f(\omega,t)d\mu\right]dt = \int\left[\int_a^b f(\omega,t)dt\right]d\mu \qquad ∎$$

关于两个积分都是 Lebesgue 积分的情形将在 1.10 节的 Fubini 定理中讨论.

习 题 1.5

1.若 $f\in L(\Omega,\mathscr{F},\mu)$，$\varepsilon>0$，证明集 $\{\omega\in\Omega:|f(\omega)|\geqslant\varepsilon\}$ 有有限测度；集 $\{\omega\in\Omega:f(\omega)\neq0\}$ 有 σ 有限测度（即为具有有限测度集的可列联）.

2.若 f 是 $\mathscr{F}$ 可测函数，$f(\omega)=0$ a.s. 证明 $f\in L(\Omega,\mathscr{F},\mu)$ 且

$$\int f d\mu = 0$$

3. 若 $f \in L(\Omega, \mathscr{F}, \mu)$，$g$ 是 Ω 上的实值可测函数，$f = g$ a.e.，证明 $g \in L(\Omega, \mathscr{F}, \mu)$，且

$$\int f d\mu = \int g d\mu$$

4. 证明若 $f \in L(\Omega, \mathscr{F}, \mu)$ 且 $\varepsilon > 0$，那么存在一个简单可测函数 φ 使得

$$\int |f - \varphi| d\mu < \varepsilon$$

5. 若 $f \in L$，g 是有界可测函数，证明其乘积 $fg \in L$.

6. 假设 $f \in L(\Omega, \mathscr{F}, \mu)$，并由它定义下列不定积分：

$$\lambda(E) = \int_E f d\mu, \quad \forall E \in \mathscr{F}$$

试证明

(1) 当且仅当 $f(\omega) \geqslant 0$ a.s. 时 $\lambda(E) \geqslant 0$，$\forall E \in \mathscr{F}$；

(2) 当且仅当 $f(\omega) = 0$ a.s. 时 $\lambda(E) = 0$，$\forall E \in \mathscr{F}$.

7. 假设 $f_1, f_2 \in L(\Omega, \mathscr{F}, \mu)$，$\lambda_1$ 和 λ_2 是由 f_1 和 f_2 分别定义的不定积分. 证明当且仅当 $f_1(\omega) = f_2(\omega)$ a.s. 时 $\lambda_1(E) = \lambda_2(E)$，$\forall E \in \mathscr{F}$.

8. 若 f 是 Ω 上的复值函数，$\mathrm{Re}f$ 和 $\mathrm{Im}f \in L(\Omega, \mathscr{F}, \mu)$，那么 f 是可积的，f 的积分定义为

$$\int f d\mu = \int \mathrm{Re}f d\mu + i \int \mathrm{Im} f d\mu$$

设 f 是复值可测函数. 证明 f 可积的充分必要条件是 $|f|$ 可积，即

$$\left| \int f d\mu \right| \leqslant \int |f| d\mu$$

[提示：若 $\int f d\mu = re^{i\theta}$，$r, \theta$ 为实数，考虑 $g(\omega) = e^{-i\theta} f(\omega)$]

9. 设 $\{f_n\}$ 是复值可测函数列，$f_n \to f$. 若存在可积函数 g 使得 $|f_n| \leqslant g$，证明

$$\int f d\mu = \lim \int f_n d\mu$$

10. 设 $\Omega = N$，$\mathscr{F}$ 是 N 所有子集所成的类，μ 是计数测度（即

$\mu(E)=E$ 中自然数个数)。 试证明当且仅当 $\sum|f(n)|$ 收敛时，$f\in L(\Omega,\mathscr{F},\mu)$，且

$$\int f d\mu=\sum_{n=1}^{\infty} f(n)$$

11. 若 $\{f_n\}$ 是 $L(\Omega,\mathscr{F},\mu)$ 中的序列，f_n 均匀收敛于 f 且 $\mu(\Omega)<+\infty$，证明

$$\int f d\mu=\lim \int f_n d\mu$$

12. 证明习题 1.5，11 缺少了条件 $\mu(\Omega)<+\infty$，命题将不再成立。

13. 设 $\Omega=R$，$\mathscr{F}=\mathscr{B}$，$f_n=nX_{[0,1/n]}$，μ 是 Lebesgue 测度。说明在 Lebesgue 控制收敛定理中 $|f_n|\leqslant g$ 是不可缺少的条件。

14. 若 $f_n\in L(\Omega,\mathscr{F},\mu)$，且

$$\sum_{n=1}^{\infty}\int|f_n|d\mu<+\infty$$

证明 $\sum f_n(\omega)\to f$ a.e.，且

$$\int f d\mu=\sum_{n=1}^{\infty}\int f_n d\mu$$

15. 假设 $f_n\in L(\Omega,\mathscr{F},\mu)$，$f_n\to f$，如果 $\lim\int|f_n-f|d\mu=0$，试证明

$$\int|f|d\mu=\lim\int|f_n|d\mu$$

16. 若 $t>0$，试证明

$$\int_0^{+\infty} e^{-tx}dx=\frac{1}{t}$$

进一步，若 $t\geqslant a>0$，那么 $e^{-tx}\leqslant e^{-ax}$。利用这里的结果和习题 1.5，12，在积分号下求导，证明

$$\int_0^{+\infty} x^n e^{-x}dx=n!$$

17. 假设 f 是 $\Omega\times[a,b]$ 到 R 的函数，$f(t,\omega)$ 是可测函数，$\forall t\in[a,b]$。并假设 $t_0, t_1\in[a,b]$，$f(\omega,t_0)$ 是 Ω 上的可积函数，$\frac{\partial f}{\partial t}(\omega,t_1)$ 存在，且存在可积函数 g 使得

$$\left|\frac{f(\omega,t)-f(\omega,t_1)}{t-t_1}\right|\leqslant g,\quad \forall\ \omega\in\Omega\quad t\in[a,b]\ t\neq t_1$$

证明

$$\left[\frac{d}{dt}\int f(\omega,t)d\mu\right]_{t=t_1}=\int\frac{\partial f}{\partial t}(\omega,t_1)d\mu$$

18. 假设 $f(\omega,t)$，对 $t\in R$ 都是可测函数($\mathscr{F}$ 可测)，$f(\omega,t)$ 是 t 的连续函数，$\forall\omega\in\Omega$。 假设存在 Ω 上的可积函数 g 和 h，使得 $|f(\omega,t)|\leqslant g(\omega)$ 且下列 Riemann 积分有

$$\int_{-\infty}^{+\infty}|f(\omega,t)|dt\leqslant h(\omega)$$

证明

$$\int_{-\infty}^{\infty}\left[\int f(\omega,t)d\mu\right]dt=\int\left[\int_{-\infty}^{+\infty}f(\omega,t)dt\right]d\mu$$

式中关于 t 的积分是 Riemann 积分。

19. 设 f 是 Ω 到 R 的可测函数，f_n 是 f 的截尾函数（参见习题1.2,6）。若 f 关于 μ 可积，证明

$$\int fd\mu=\lim\int f_nd\mu$$

反之，若

$$\sup\int|f_n|d\mu<+\infty$$

证明 f 可积。

1.6 L_p 空间——Banach 空间

提要 本节首先引入一般线性空间上的范数概念，然后在证明 Hölder 不等式，导出 Minkowski 不等式之后，引入 L_p 空间上的范数。最后证明 L_p 空间在范数下完备以及 L_∞ 的完备性。

这里先讲述一般线性空间。

设 V 是一个集，假定在 V 中规定线性运算——元素的加法运算以及实数(或复数)与 V 中元素的乘法运算，且满足下列条件：

(1) V 关于加法成交换群，即对任何 $u, v\in V$，都存在 $w\in V$，记作 $w=u+v$，我们称 w 是 u,v 的和，这一运算有以下性质：

(i) $u+v=v+u$;

(ii) $(u+v)+w=u+(v+w)$;

(iii) V 中存在唯一的元素 0 (称为零元素)，使得对于任何 $u\in V$，$0+u=u$ 成立；

(iv) 对 V 中任一元素 u，均存在唯一元素 u' 使 $u+u'=0$，u' 记作 $-u$。

(2) 对任何 $u\in V$ 以及实数 a (或复数)，存在 $au\in V$，我们称 au 是 u 与 a 之数积。数积运算具有如下性质：

(i) $1\cdot u=u$;

(ii) $a(bu)=(ab)u$，a,b 是实数(或复数)；

(iii) $(a+b)u=au+bu$，$a(u+v)=au+av$；其中 a,b 是实数(或复数)，$u,v\in V$。那么 V 称为线性空间。

容易验证，测度空间 $(\Omega, \mathscr{F}, \mu)$ 上的全体可积函数组成一个线性空间，在这个空间上定义范数后具有 Banach 空间结构。

定义 1.6.1 设 V 是一实线性空间(向量空间)，V 上的实函数 N 称为范数，如果 N 满足下列条件：

(1) $N(v)\geqslant 0$，$\forall\ v\in V$;

(2) 当且仅当 $v=0$ 时，$N(v)=0$;

(3) $N(\alpha v)=|\alpha|N(v)$，$\forall\ v\in V$ 和实数 α;

(4) $N(u+v)\leqslant N(u)+N(v)$，$\forall u,v\in V$。

如果 N 满足条件(1),(3),(4)，那么 N 称为 V 上的半范数

或拟范数。线性空间V与V上的范数N组成线性赋范空间。

例1

(i) 实数的绝对值是R的范数；

(ii) n维实向量组成的线性空间，其范数可以定义为

$$N_1(u_1,\cdots,u_n)=|u_1|+\cdots+|u_n|$$

$$N_p(u_1,\cdots,u_n)=\{|u_1|^p+\cdots+|u_n|^p\}^{1/p},\quad p\geqslant 1$$

$$N_\infty(u_1,\cdots,u_n)=\sup\{|u_1|,\cdots,|u_n|\}$$

容易验证N_1和N_∞都是范数，N_p满足条件(1)，(2)，(3)。利用 Minkowski 不等式可以证明N_p也满足条件(4)。

(iii) l_1是使$N_1(u)=\sum|u_n|<+\infty$的全体无穷序列$u=\{u_n\}$组成的线性空间。它在范数N_1下是一个线性赋范空间。

类似地，l_p是使$N_p(u)=\{\sum|u_n|^p\}^{1/p}<+\infty$，$1\leqslant p<+\infty$的全体无穷序列$u=\{u_n\}$组成的线性空间。它是在范数$N_p$下的线性赋范空间。

(iv) 定义在无限集Ω上的全体实值函数是不能赋范的。然而Ω上的全体有界实函数是可以赋范的，其范数为

$$N(f)=\sup\{|f(\omega)|:\omega\in\Omega\}$$

特别地，$[a,b]$上连续函数组成的线性空间是可以赋范的。

上面都是线性空间上的范数的例子，下面给出一些半范数(或拟范数)的例子。

例2

(i) 在空间R^n上，考虑半范数：

$$N_0(u_1,\cdots,u_n)=\sup\{|u_2|,\cdots,|u_n|\}$$

这里当且仅当$u_2=\cdots=u_n=0$时，$N_0(u)=0$，但$u\neq 0$。

(ii) $[0,1]$上的全体连续函数组成线性空间$C[0,1]$，定义半范数

$$N_0(f)=\sup_{0\leqslant x\leqslant \frac{1}{2}}|f(x)|$$

这里当且仅当 $f(x)$ 在闭区间$\left[0,\frac{1}{2}\right]$上等于零时 $N_0(f)=0$.

(iii) $[a,b]$到R的具有连续导函数的函数组成一线性空间，f 的半范数定义为

$$N_0(f)=\sup_{a\leqslant x\leqslant b}|f'(x)|$$

当且仅当 f 在$[a,b]$上为常数时 $N_0(f)=0$.

定义 1.6.2 设 $(\Omega,\mathscr{F},\mu)$ 是测度空间，$f\in L(\Omega,\mathscr{F},\mu)$. 定义

$$N_\mu(f)=\int|f|d\mu$$

可以证明$N_\mu(f)$是 $L(\Omega,\mathscr{F},\mu)$ 上的半范数.

引理 1.6.1 定义下列运算：

$$(f+g)(\omega)=f(\omega)+g(\omega),\quad \forall f,g\in L$$

$$(\alpha f)(\omega)=\alpha f(\omega),\quad \forall f\in L$$

$L(\Omega,\mathscr{F},\mu)$ 关于这些运算是一个线性空间. $N_\mu(f)$ 是半范数，当且仅当 $f(\omega)=0$ a.e. 时 $N_\mu(f)=0$.

证明 由定理 1.5.2 可知 $L(\Omega,\mathscr{F},\mu)$ 是线性空间. 显然有 $N_\mu(f)\geqslant 0$, $\forall f\in L$.

$$N_\mu(\alpha f)=\int|\alpha f|d\mu=|\alpha|\int|f|d\mu=|\alpha|N_\mu(f)$$

利用三角不等式得

$$N_\mu(f+g)=\int|f+g|d\mu\leqslant\int(|f|+|g|)d\mu$$

$$=\int|f|d\mu+\int|g|d\mu=N_\mu(f)+N_\mu(g)$$

因此N_μ是L上的半范数. 由推论 1.4.3 可得：当且仅当 $f(\omega)=0$, a.s. 时，$N_\mu(f)=0$. ▮

如果把两个几乎处处相等的函数看成同一函数，即我们用函数等价类代替函数，那么 $N_p(f)$ 是范数，$L(\Omega, \mathscr{F}, \mu)$ 是线性赋范空间，把这个意思写成严格的定义形式：

定义 1.6.3 若 $L = L(\Omega, \mathscr{F}, \mu)$ 中的两个函数几乎处处相等，则称它们关于 (μ) 等价. $[f]$表示 L 中所有与 f 等价的函数. L 中的一切等价类组成的空间记作 $L_1 = L_1(\Omega, \mathscr{F}, \mu)$. 若 $[f] \in L$，定义范数

$$\|[f]\|_1 = \int |f| d\mu \tag{1.6.1}$$

定理 1.6.1 $L_1(\Omega, \mathscr{F}, \mu)$ 是线性赋范空间.

证明 我们认为在 L_1 中已经定义了下列运算：

$$\alpha[f] = [\alpha f], \quad [f] + [g] = [f + g]$$

L_1 中的零元素是 $[0]$. 现在我们只需验证式 (1.6.1) 定义了 L_1 上的范数. 显然 $\|[f]\|_1 \geq 0$. 且 $\|[0]\|_1 = 0$.

若 $\|[f]_1\|_1 = 0$，那么

$$\int |f| d\mu = 0$$

因此 $f(\omega) = 0$, a.s.，于是 $[f] = [0]$. 容易看到 $\|[f]\|_1$ 满足定义 1.6.1 的条件(3)和(4). 所以 $\|\cdot\|_1$ 是 L_1 上的范数. ∎

我们应该记住 L_1 是以 L 中等价类作为元素的，记住了这一点，我们把 $\|[f]\|_1$ 记作 $\|f\|_1$.

定义 1.6.4 若 $1 \leqslant p < +\infty$，$L_p = L_p(\Omega, \mathscr{F}, \mu)$ 包含了使得

$$\int |f|^p d\mu < +\infty \tag{1.6.2}$$

的可测函数等价类组成的空间. 两个函数等价意思指两个函数几乎处处相等. 定义

$$\|f\|_p = \left\{\int |f|^p d\mu\right\}^{1/p} \tag{1.6.3}$$

当 $p=1$ 时，L_p 退化为 L_1 的情形。

下面我们将证明当 $1 \leqslant p < +\infty$ 时，L_p 在式(1.6.3)定义的范数下是完备的线性赋范空间。因此 L_p 是 Banach 空间。

为了证明由式(1.6.3)定义的函数确实是范数，我们先证明下列基本不等式。

定理 1.6.2 (Hölder 不等式) 若 $f \in L_p$，$g \in L_q$，这里 $p>1$ 且 $\left(\frac{1}{p}+\frac{1}{q}\right)=1$，那么 $fg \in L_1$，且

$$\|fg\|_1 \leqslant \|f\|_p \|g\|_q$$

证明 设 α 是实数，$0<\alpha<1$。φ 是定义在 $[0,+\infty)$ 上的函数

$$\varphi(t)=\alpha t-t^{\alpha}$$

容易验证：当 $0<t<1$ 时，$\varphi'(t)<0$，当 $t>1$ 时

$$\varphi'(t)>0$$

由中值定理可知 $\varphi(t) \geqslant \varphi(1)$。当且仅当 $t=1$ 时 $\varphi(t)=\varphi(1)$。所以

$$t^{\alpha} \leqslant \alpha t+(1-\alpha), \quad \forall t \geqslant 0$$

若 $a \geqslant 0, b \geqslant 0$，令 $t=a/b$ 可得

$$a^{\alpha} b^{1-\alpha} \leqslant \alpha a+(1-\alpha) b$$

当且仅当 $a=b$ 时等号成立。

p 和 q 满足 $p<1$ 且 $\frac{1}{p}+\frac{1}{q}=1$。取 $\alpha=1/p$，若 A，B 为任何非负实数 ($A^p=a, B^q=b$)，那么

$$AB \leqslant \frac{A^p}{p}+\frac{B^q}{q} \tag{1.6.4}$$

当且仅当 $A^p=B^q$ 时式(1.6.4)等号成立。

假设 $f \in L_p$，$g \in L_q$ 且 $\|f\|_p \neq 0$，$\|g\|_q \neq 0$ 时，fg 可测。令 $A=|f(\omega)|/\|f\|_p$，$B=|g(\omega)|/\|g\|_q$，并代入式

(1.6.4)得

$$\frac{|f(\omega)g(\omega)|}{\|f\|_p\|g\|_q} \leqslant \frac{|f(\omega)|^p}{p\|f\|_p^p} + \frac{|g(\omega)|^q}{q\|g\|_q^q}$$

因为右边的两项均可积，由推论 1.5.1 和定理 1.5.2 可知 fg 可积。上式两边积分后得

$$\frac{\|fg\|_1}{\|f\|_p\|g\|_q} \leqslant \frac{1}{p} + \frac{1}{q} = 1$$

因此 Hölder 不等式成立。 ▮

Hölder 不等式告诉我们：当 $p > 1$，$\frac{1}{p} + \frac{1}{q} = 1$，$f \in L_p$，$g \in L_q$ 时，fg 是可积的。

两个实数 p，q，若满足 $p > 1$，$\frac{1}{p} + \frac{1}{q} = 1$ 或等价地 $p + q = pq$，则称 p, q 是共轭指标。

当 $p = 2$ 时 $q = 2$，称为自共轭指标。由此可知两个 L_2 中函数之积是可积的。

定理 1.6.3 (Cauchy-Bunyakovskii-Schwarz不等式) 若 $f, g \in L_2$，那么 fg 可积且

$$\left|\int fg d\mu\right| \leqslant \int |fg| d\mu \leqslant \|f\|_2\|g\|_2 \qquad (1.6.5)$$

证明 这仅仅是 Hölder 不等式 $p = 2$ 的特例。 ▮

定理 1.6.4 (Minkowski 不等式) 若 f 和 $h \in L_p$，$p \geqslant 1$，那么 $f + h \in L_p$ 且

$$\|f + h\|_p \leqslant \|f\|_p + \|h\|_p \qquad (1.6.6)$$

证明 $p = 1$ 已经证明，因此假定 $p > 1$。显然 $f + h$ 可测。因为

$$\begin{aligned}|f + h|^p &\leqslant [2\sup\{|f|, |h|\}]^p \\ &\leqslant 2^p\{|f|^p + |h|^p\}\end{aligned}$$

由推论 1.5.1 和定理 1.5.2 可知 $f + h \in L_p$。更进一步

$$|f+h|^p = |f+h||f+h|^{p-1}$$
$$\leqslant |f||f+h|^{p-1} + |h||f+h|^{p-1} \quad (1.6.7)$$

因为 $f+h\in L_p$，那么$|f+h|^p\in L_1$。又因为 $p=(p-1)q$，因此 $|f+h|^{p-1}\in L_q$。利用 Hölder 不等式可得

$$\int |f||f+h|^{p-1}d\mu \leqslant \|f\|_p \left\{\int |f+h|^{(p-1)q}d\mu\right\}^{1/q}$$
$$= \|f\|_p\|f+h\|_p^{p/q}$$

式(1.6.7)右端第二项作类似处理得

$$\|f+h\|_p^p \leqslant \|f\|_p\|f+h\|_p^{p/q} + \|h\|_p\|f+h\|_p^{p/q}$$
$$= \{\|f\|_p + \|h\|_p\}\|f+h\|_p^{p/q}$$

若 $\|f+h\|_p=0$，式(1.6.6)是平凡等式。若 $\|f+h\|_p\neq 0$不等式两边同除 $\|f+h\|_p^{p/q}$ 得

$$\|f+h\|_p^{p-p/q} \leqslant \|f\|_p + \|h\|_p$$

因为 $p-\frac{p}{q}=1$，Minkowshi 不等式得证。 ∎

容易验证 L_p 在由式(1.6.3)定义的函数中满足 (1),(2),(3)条件，利用 Minkowski 不等式可以证明它也满足定义中条件(4)。

L_p 在范数式(1.6.3)下完备，那么什么是完备呢?

定义 1.6.5 L_p 中序列 $\{f_n\}$ 称为 Cauchy 列，如果对任何 $\varepsilon>0$，存在正整数 $M(\varepsilon)$，当 $m,n\geqslant M(\varepsilon)$时，$\|f_m-f_n\|_p<\varepsilon$。若 $\{f_n\}$ 是 L_p中的序列且 $f\in L_p$，对任何 $\varepsilon>0$,存在正整数 $N(\varepsilon)$，当 $n\geqslant N(\varepsilon)$ 时，$\|f_n-f\|_p<\varepsilon$，则称 $\{f_n\}$ 以范数收敛到 L_p 中的元素 f。若线性赋范空间中每一个 Cauchy 的极限都属于本空间，则称该空间是完备的。

引理 1.6.2 若 $\{f_n\}$ 是一个收敛于 L_p 中 f 的序列，那么它必为 Cauchy 列。

证明 若 $m, n\geqslant N(\varepsilon/2)$，那么

$$\|f-f_m\|_p<\varepsilon/2,\ \|f-f_n\|<\varepsilon/2$$

因此

$$\|f_m-f_n\|_p\leqslant\|f_m-f\|_p+\|f_n-f\|_p<\varepsilon$$

即 $\{f_n\}$ 为 Cauchy 列。 ∎

现在我们来证明 L_p 中的 Cauchy 列收敛到 L_p 中的元素。这一结果有时称为 Riesz-Fischer 定理。

定理 1.6.5(完备性定理) 若 $1\leqslant p<+\infty$，那么 L_p 在范数

$$\|f\|_p=\left\{\int|f|^p d\mu\right\}^{1/p}$$

下是线性赋范完备空间(即 Banach 空间)。

证明 上面已经证明了 L_p 是线性赋范空间。为了建立 L_p 的完备性，设 $\{f_n\}$ 是 L_p 中任一 Cauchy 序列。因此对 $\varepsilon>0$，存在 $M(\varepsilon)$ 使得 $m,n\geqslant M(\varepsilon)$ 时

$$\int|f_m-f_n|^p d\mu=\|f_m-f_n\|_p^p<\varepsilon^p \tag{1.6.8}$$

设 $\{f_n\}$ 存在一个子序列 $\{g_k\}$

$$\|g_{k+1}-g_k\|_p<2^{-k},\quad\forall k\in N$$

定义

$$g(\omega)=|g_1(\omega)|+\sum_{k=1}^{\infty}|g_{k+1}(\omega)-g_k(\omega)| \tag{1.6.9}$$

$g\in M^+(\Omega,\mathscr{F})$。由 Fatou 引理可知

$$\int|g|^p d\mu\leqslant\liminf_{n\to\infty}\int\left\{|g_1|+\sum_{k=1}^{n}|g_{k+1}-g_k|\right\}^p d\mu$$

两边开 p 次方，并利用 Minkowski 不等式可得

$$\left\{\int|g|^p d\mu\right\}^{1/p}\leqslant\liminf_{n\to\infty}\left\{\|g_1\|_p+\sum_{k=1}^{n}\|g_{k+1}-g_k\|_p\right\}$$
$$\leqslant\|g_1\|_p+1$$

令 $E=\{\omega\in\Omega: g(\omega)<+\infty\}, E\in\mathscr{F}$，$\mu(\Omega-E)=0$。因此 $g\in L_p$，式（1.6.9）中的级数几乎处处收敛，且 $g\chi_E\in L_p$。

现在定义

$$f=\begin{cases} g_1(\omega)+\sum_{k=1}^{\infty}\{g_{k+1}(\omega)-g_k(\omega)\}, & \forall\omega\in E \\ 0, & \forall\omega\notin E\end{cases}$$

因为 $|g_k|+\sum_{j=k}^{\infty}|g_{j+1}-g_j|\leqslant g$，且 $g_k\to f$ a.e.，利用控制收敛定理可知 $f\in L_p$。

又因为 $|f-g_k|^p\leqslant 2^p g^p$，由控制收敛定理可知 $\lim\|f-g\|_p=0$，因此 $g_k\xrightarrow{p\text{范}}f\in L_p$

根据式(1.6.8)，对充分大的 k 有

$$\int|f_m-g_k|^p d\mu<\varepsilon^p$$

利用 Fatou 引理，当 $m\geqslant M(\varepsilon)$ 时，可得

$$\int|f_m-f|^p d\mu\leqslant\liminf_{k\to\infty}\int|f_m-g_k|^p d\mu\leqslant\varepsilon^p$$

这就证明了 $\{f_n\}$ 在范数下收敛到 L_p 的元素。 ▮

完备线性赋范空间称为 Banach 空间。因此 L_p 是一 Banach 空间。

定义 1.6.6 $L_\infty=L_\infty(\Omega,\mathscr{F},\mu)$ 是几乎处处有界的等价类函数的集合。若 $f\in L_\infty, N\in\mathscr{F}$ 且 $\mu(N)=0$，我们定义

$$S(N)=\sup_{\omega\notin N}\{|f(\omega)|\}$$

且

$$\|f\|_\infty=\inf\{S(N): N\in\mathscr{F}, \mu(N)=0\} \quad (1.6.10)$$

L_∞ 中的元素称为本质有界函数。

容易验证式(1.6.10)定义了范数。

定理 1.6.6 在式(1.6.10)的范数意义下，L_∞ 是完备的线性赋范空间。

证明 显然 L_∞ 是一线性空间且 $\|f\|_\infty \geqslant 0$，$\|0\|_\infty = 0$，$\|\alpha f\|_\infty = |\alpha| \|f\|_\infty$；若 $\|f\|_\infty = 0$，那么存在 $N_k \in \mathscr{F}$，$\mu(N_k) = 0$ 使得

$$|f(\omega)| \leqslant 1/k, \quad \forall \omega \notin N_k$$

令 $N = \bigcup_{k=1}^{\infty} N_k$，那么 $N \in \mathscr{F}$

且

$$\mu(N) = 0, \quad |f(\omega)| = 0, \quad \forall \omega \notin N$$

所以 $f(\omega) = 0$ a.e. [这说明 $\|f\|_\infty$ 满足定义 1.6.1 的条件(2)]。

若 $f, g \in L_\infty$，存在 $N_1, N_2 \in \mathscr{F}$，且 $\mu(N_1) = \mu(N_2) = 0$，使得

$$|f(\omega)| \leqslant \|f\|_\infty, \quad \forall \omega \notin N_1$$

$$|g(\omega)| \leqslant \|g\|_\infty, \quad \forall \omega \notin N_2$$

于是 $|f(\omega) + g(\omega)| \leqslant \|f\|_\infty + \|g\|_\infty$，$\forall \omega \notin (N_1 \cup N_2)$。因此

$$\|f + g\|_\infty \leqslant \|f\|_\infty + \|g\|_\infty$$

满足定义 1.6.1 条件(4)。

下面证明 L_∞ 的完备性。设 $\{f_n\}$ 是 L_∞ 中的 Cauchy 列，$M \in \mathscr{F}$ 且 $\mu(M) = 0$ 使得

$$|f_n| \leqslant \|f_n\|_\infty, \quad \forall \omega \notin M, \quad n = 1, 2, \cdots$$

且

$$|f_n - f_m| \leqslant \|f_n - f_m\|_\infty, \quad \forall \omega \notin M \qquad n, m = 1, 2, \cdots$$

因此 $\{f_n\}$ 在 $\Omega - M$ 上均匀收敛，令

$$f(\omega)=\begin{cases}\lim f_n(\omega), & \omega \notin M\\ 0, & \omega\in M\end{cases}$$

由此可知 f 可测，容易验证 $\|f_n-f\|_\infty\to 0$。因此 L_∞ 是完备的。 ▮

习　　题　　1.6

1. $C[0,1]$ 表示$[0,1]$上连续函数组成的线性空间。定义

$$N_0(f)=|f(0)|$$

证明 $N_0(f)$ 在 $C[0,1]$ 上是半范数。

2. $C[0,1]$ 定义如习题1.6,1

$$N_1(f)=\int_0^1 |f(x)|dx$$

证明 N_1 是 $C[0,1]$ 上的半范数。

定义，$n\geqslant 1$

$$f_n=\begin{cases}0, & 0\leqslant x\leqslant\left(1-\frac{1}{n}\right)/2\\ ax+b, & \left(1-\frac{1}{n}\right)/2<x\leqslant\frac{1}{2}\\ 1, & \frac{1}{2}<x\leqslant 1\end{cases}$$

证明 $\{f_n\}$ 是一 Cauchy 序列。但在 N_1 它不收敛到 $C[0,1]$ 中的元素。

3.设 N 是线性空间 V 上的范数，d 定义如下：

$$d(u,v)=N(u-v).\ \forall u,v\in V$$

试证明 d 是 V 上的距离，即

(1) $d(u,v)\geqslant 0$，$\forall\ u,v\in V$；

(2) $d(u,v)=0$，当且仅当 $u=v$；

(3) $d(u,v)=d(v,u)$；

(4) $d(u,v)\leqslant d(u,w)+d(w,v)$。

4.若 $f\in L_1(\Omega,\mathscr{F},\mu)$，$\varepsilon>0$ 存在一简单可测函数 φ 使得 $\|f-\varphi\|_1<\varepsilon$。把这一结论推广到 L_p 空间。对于 L_∞ 这一命题是否成立？

5. 若 $f \in L_p$，$1 \leqslant p < \infty$，$E = \{\omega \in \Omega: |f(\omega)| \neq 0\}$ 试证明 E 是 σ 有限集.

6. 若 $f \in L_p, E_n = \{\omega \in \Omega: |f(\omega)| \geqslant n\}$，试证明当 $n \to \infty$ 时 $\mu(E_n) \to 0$。

7. 设 $\Omega = N$（全体自然数），μ 是计数测度. N 上的函数 $f(n) = \frac{1}{n}$，试证明 $f \notin L_1$ 但 $f \in L_p$，这里 $1 < p \leqslant \infty$。

8. 设 $\Omega = N, \mathscr{F}$ 是 N 一切子集所成的类

$$\lambda(E) = \sum_{n \in E} \frac{1}{n^2}, \quad E \in \mathscr{F}$$

证明 $\mu(\Omega) < +\infty$. f 是 N 上的函数 $f(n) = \sqrt{n}$. 证明当且仅当 $1 \leqslant p < 2$ 时 $f \in L_p$.

9. 设 $(\Omega, \mathscr{F}, \mu)$ 是有限测度空间. f 是 $\mathscr{F}$ 可测函数，$E_n = \{\omega \in \Omega: (n-1) \leqslant |f(\omega)| < n\}$. 试证明当且仅当

$$\sum_{n=1}^{\infty} n\mu(E_n) < +\infty$$

时，$f \in L_1$.

更进一步，当且仅当

$$\sum_{n=1}^{\infty} n^p \mu(E_n) < +\infty$$

时，$f \in L_p$，其中 $1 \leqslant p < \infty$.

10. 设 $(\Omega, \mathscr{F}, \mu)$ 是有限测度空间，$f \in L_p$，试证明 $f \in L_r$，其中 $1 \leqslant r \leqslant p$.

[提示：利用习题 1.6,9 或不等式 $|f|^r \leqslant 1 + |f|^p, |f|^r \in L_{p/r}, g = 1$，再利用 Hölder 不等式，可得 $\|f\|_r \leqslant \|f\|_p \mu(\Omega)^s$，其中 $s = \frac{1}{r} - \frac{1}{p}$. 若 $\mu(\Omega) = 1$，那么 $\|f\|_r \leqslant \|f\|_p$].

11. 设 $\Omega = (0, +\infty)$，μ 是 Ω 上 Lebesgue 测度，令 $f(x) = x^{-1/2} \times (1 + |\log x|)^{-1}$，试证明当且仅当 $p = 2$ 时 $f \in L_p$.

12. 设 $(\Omega, \mathscr{F}, \mu)$ 是测度空间，$f \in L_{p_1}$ 且 $f \in L_{p_2}$，这里 $1 \leqslant p_1 < p_2 < +\infty$. 证明对于任何 $p: p_1 \leqslant p \leqslant p_2$ 有 $f \in L_p$.

13. 设 $f\in L_p(\Omega,\mathscr{F},\mu)$，$1\le p<\infty$ 且 $\varepsilon\ge 0$． 证明存在 $E_\varepsilon\in\mathscr{F}$，$\mu(E_\varepsilon)<+\infty$，而且 $F\cap E_\varepsilon=\varnothing$ $\forall$ $F\in\mathscr{F}$，那么 $\|f\chi_F\|_p<\varepsilon$．

14. 设 $f_n\in L_p(\Omega,\mathscr{F},\mu)$，$1\le p<\infty$，$\beta_n$ 定义为

$$\beta_n(E)=\left\{\int_E|f_n|^p d\mu\right\}^{1/p}$$

证明 $|\beta_n(E)-\beta_m(E)|\le\|f_n-f_m\|_p$．若 $\{f_n\}$ 是 L_p 中的 Cauchy 列，证明对每一个 $E\in\mathscr{F}$，$\lim\beta_n(E)$ 存在．

15. 设 f_n,β_n 由习题1.6,14给出。若 $\{f_n\}$ 是 Cauchy 列，$\varepsilon>0$，试证明存在 $E_\varepsilon\in\mathscr{F}$，$\mu(E_\varepsilon)<+\infty$ 使得 $F\cap E_\varepsilon=\varnothing, F\in\mathscr{F}$ 且

$$\beta_n(F)<\varepsilon,\ \forall n\in N$$

16. 设 f_n，β_n 由习题1.6，15给出。假设 $\{f_n\}$ 是一 Cauchy 序列．若 $\varepsilon>0$，则存在 $\delta(\varepsilon)>0$ $\mu(E)<\delta(\varepsilon)$，$E\in\mathscr{F}$． 证明 $\beta_n(E)<\varepsilon$，$\forall n\in N$ 成立．

17. 若 $f\in L_p$，$1\le p\le\infty$ 且 $g\in L_\infty$ 试证明 $fg\in L_p$ 且

$$\|fg\|_p\le\|f\|_p\|g\|_\infty.$$

18. 当且仅当 $\mu(\Omega)<+\infty$ 时 $L_\infty\subset L_1$．若 $\mu(\Omega)=1$ 且 $f\in L_\infty$，试证明

$$\|f\|_\infty=\lim_{p\to\infty}\|f\|_p.$$

1.7 各种收敛关系

提要 本节讨论了几乎处处收敛，几乎一致（均匀）收敛，依测度收敛和平均收敛等重要概念以及它们之间的相互关系。这些概念在第二、三章概率论中起着重要的作用。本节还证明了 Riesz、Egoroff 和 Vitali 等人命名的重要定理。

可测函数的各种收敛关系在控制论、信号处理以及其他有关学科的文献中经常出现。理解并搞清各种收敛性之间的

关系成了学习测度论的重要课题之一。

本节我们只考虑在已知测度空间$(\Omega,\mathscr{F},\mu)$上的实可测函数和 L_p 空间，$1\leqslant p<\infty$。当然在有些实际应用中还需要广义实可测函数或 L_∞空间，这些只要在本节讨论的基础上作相应的修正，是不难获得相应结果的。

在前面几节中，我们已经提到过四种收敛关系：点点收敛，几乎处处收敛，均匀收敛以及 L_p 中的收敛。为了讨论的方便，这里重新叙述一下这些收敛关系的定义。

可测函数列 $\{f_n\}$ 均匀(或一致)收敛于 f 是指：若对 $\varepsilon>0$ 存在自然数 $N(\varepsilon)$，且当 $n\geqslant N(\varepsilon)$，$\omega\in\Omega$ 时有

$$|f_n(\omega)-f(\omega)|<\varepsilon$$

可测函数到$\{f_n\}$点点收敛于 f 是指：对 $\varepsilon>0$ 和 $\omega\in\Omega$，存在自然数 $N(\varepsilon,\omega)$，且当 $n\geqslant N(\varepsilon,\omega)$ 时有

$$|f_n(\omega)-f(\omega)|<\varepsilon$$

可测函数列 $\{f_n\}$ 几乎处处收敛是指：存在集 $M\in\mathscr{F}$，$\mu(M)=0$。对于任何 $\varepsilon>0$，$\omega\in\Omega-M$，存在自然数 $N(\varepsilon,\omega)$ 使得在 $n\geqslant N(\varepsilon,\omega)$ 时有 $|f_n-f|<\varepsilon$。

显然，均匀收敛的函数列必定点点收敛；点点收敛必几乎处处收敛。容易证明这些命题的逆命题是不成立的。(当然，如果 Ω 是有限个元素组成的集合，那么点点收敛与均匀收敛是一样的；如果测度为零的集必为空集，那么几乎处处收敛与点点收敛是一样的。除了这些特殊情形外，一般地说，逆命题不成立)

这里我们进一步考虑上一节引进的 L_p 收敛性概念。L_p 是表示 p 次可积的等价类，$1\leqslant p<+\infty$。

定义 1.7.1 $L_p=L_p(\Omega,\mathscr{F},\mu)$ 中的序列，$\{f_n\}$收敛于 $f\in L_p$，如果对任给的 $\varepsilon>0$，存在自然数 $N(\varepsilon)$，当 $n\geqslant N(\varepsilon)$ 时

$$\|f_n - f\|_p = \left\{\int |f_n - f|^p d\mu\right\}^{1/p} < \varepsilon, \quad p \geq 1$$

有时我们把这种收敛关系称为(p 次)平均收敛.

$\{f_n\}$ 称为 L_p 中的 Cauchy 列,如果对任给的 $\varepsilon > 0$, 存在自然数 $N(\varepsilon)$, 当 $m, n > N(\varepsilon)$ 时,有

$$\|f_m - f_n\| = \left\{\int |f_m - f_n|^p d\mu\right\}^{1/p} < \varepsilon$$

根据定理 1.6.5,若 $\{f_n\}$ 是 L_p 中的 Cauchy 列,那么必存在 $f \in L_p$, f_n 平均收敛于 f. 这是因为 L_p 在平均收敛意义下是完备的.

L_p 在平均收敛意义下完备,即关于平均收敛的极限运算封闭, 然而在其他极限意义下就不一定封闭. 例如, 存在 $\{f_n\} \subset L_p$ 且 f_n 均匀收敛于 f, $f \notin L_p$ (见习题 1.7,2). 如果 $\mu(\Omega) < +\infty$, 这种情况就不会发生.

定理 1.7.1 假设 $\mu(\Omega) < +\infty, f_n \in L_p, n = 1,2,\cdots$, f_n 均匀收敛于 f,那么 $f \in L_p$ 且 $\{f_n\}$ 平均收敛于 f.

证明 由于 f_n 均匀收敛于 f, 因此对任给 $\varepsilon > 0$, 存在 $N(\varepsilon)$, 当 $n \geq N(\varepsilon)$ 时

$$|f_n - f| < \varepsilon, \quad \forall\ \omega \in \Omega$$

于是 $|f_n - f|^p < \varepsilon^p$, $|f_n - f|$ p 次可积且

$$\|f_n - f\|^p = \left\{\int |f_n - f|^p d\mu\right\}^{1/p} \tag{1.7.1}$$

$$\leq \left\{\int \varepsilon^p d\mu\right\}^{1/p} = \varepsilon \mu(\Omega)^{1/p}$$

因此 f_n 平均收敛于 f,且 $f \in L_p$. ∎

注意,若 $\mu(\Omega) < +\infty$, $f_n \in L_p$, $n = 1,2,\cdots$,且 f_n 点点收敛于 f,那么 f 不一定属于 L_p. 然而, 如果 $\mu(\Omega) < \infty$ 再加上存在 $g \in L_p$,使得

$$|f_n| \leq g, \quad \forall \omega \in \Omega, \quad n \in N$$

那么 f_n 平均收敛于 $f\in L_p$ 成立.

定理 1.7.2 设 $f_n\in L_p$, $n=1,2,\cdots$, 且 f_n 几乎处处收敛于 f. 若存在 $g\in L_p$ 使得

$$|f_n|\leqslant g,\ \forall\omega\in\Omega,\ n\in N \tag{1.7.2}$$

那么 $f\in L_p$ 且 f_n 平均收敛于 f.

证明 由不等式(1.7.2)可知, $|f|\leqslant g$, a.e.. 因为 $g\in L_p$, 由推论 1.5.1 可得 $f\in L_p$.
又因为

$$|f_n-f|^p\leqslant[2g]^p,\ \text{a.e.}$$

且 $\lim|f_n-f|^p=0$, a.e., $2^pg^p\in L_1$, 由 Lebesgue 控制收敛定理,有

$$\lim\int|f_n-f|^pd\mu=0$$

即 f_n 平均收敛于 f. ∎

推论 1.7.1 若 $\mu(\Omega)<+\infty$, $\{f_n\}$ 是 L_p 中的序列, $\lim f_n=f$, a.e.. 又存在常数K使得

$$|f_n(\omega)|\leqslant K,\ \forall\omega\in\Omega,\ n\in N \tag{1.7.3}$$

那么 $f\in L_p$, f_n 平均收敛于 f.

证明 因为 $\mu(\Omega)<+\infty$, 常数函数属于 L_p,由定理 1.7.2 可得此推论. ∎

可能人们会认为平均收敛也可以推出几乎处处收敛关系,然而这是不成立的. 我们可以给出这样的例子: f_n 平均收敛于 f,然而 f_n 在 Ω 中任何点上都不收敛于 f.

例 设 $\Omega=[0,1]$, $\mathscr{B}$ 是 Lebesgue 可测集类, λ 是 $\mathscr{B}$ 上 Lebesgue 测度. 考虑区间序列 $[0,1],[0,1/2],[1/2,1],[0,1/3],[1/3,2/3],[2/3,1],[0,1/4],[1/4,1/2],[1/2,3/4],[3/4,1],[0,1/5],[1/5,2/5],\cdots$. f_n 表示区间序列中第 n 个区间的示性函数, $f=0$. 若 $n\geqslant\frac{m(m+1)}{2}$ $(=1+$

$2+\cdots+m)$, $\lambda(E_n)\leqslant\frac{1}{m}$. 于是

$$\|f_n-f\|_p^p=\int|f_n-f|^p d\lambda$$

$$=\int f_n d\lambda\leqslant 1/m$$

f_n 平均收敛于 f. 然而,对任何一点 $\omega\in\Omega$, $f_n(\omega)$ 有一个全是 1 的子序列,一个全是零的子序列. 因此 f_n 在 $[0,1]$ 中任何一点都不收敛.

虽然平均收敛不能推出几乎处处收敛，然而能推出另一类重要收敛关系.

定义 1.7.2 实可测函数列 $\{f_n\}$ 依测度收敛于可测函数 f 是指：如果对任给 $\varepsilon>0$, 有

$$\lim_{n\to\infty}\mu(\{\omega\in\Omega:|f_n(\omega)-f(\omega)|\geqslant\varepsilon\})=0$$

若

$$\lim_{n,m\to\infty}\mu(\{\omega\in\Omega:|f_m(\omega)-f_n(\omega)|\geqslant\varepsilon\})=0$$

则称 $\{f_n\}$ 是依测度收敛的 Cauchy 列.

若 f_n 均匀收敛于 f,那么集

$$\{\omega\in\Omega:\ |f_n(\omega)-f(\omega)|\geqslant\varepsilon\}$$

对于充分大的 n 是空集. 因此均匀收敛包含了依测度收敛.

一般地说，点点收敛(或几乎处处收敛)并不包含依测度收敛. 但如果 $\mu(\Omega)<+\infty$, 那么由几乎处处收敛可推出依测度收敛(这就是著名的 Egoroff 定理).

容易看到 (p 次) 平均收敛包含依测度收敛. 事实上,如果对任给 $\alpha>0$, 有

$$E_n(\alpha)=\{\omega\in\Omega:|f_n(\omega)-f(\omega)|\geqslant\alpha\}$$

那么

$$\int |f_n - f|^p d\mu \geqslant \int_{E_n(\alpha)} |f_n - f|^p d\mu$$
$$\geqslant d^p \mu(E_n(\alpha))$$

因为 f_n 平均收敛，$\|f_n - f\|_p \to 0$，因此

$$\lim_{n\to\infty} \mu(E_n(\alpha)) = 0$$

即 $f_n \xrightarrow{\mu} f$（f_n 依测度收敛于 f）。

从上例我们看到一个测度收敛的序列，可以在任何点上都不收敛。

然而 F. Riesz 已经证明：若 f_n 依测度收敛于 f，那么必存在一个子序列 $f_{n_k} \to f$，a.e.。 ▮

定理 1.7.3 若 $\{f_n\}$ 是依测度收敛的 Cauchy 列，那么必存在子序列 $\{g_n\}$ 几乎处处且依测度收敛于可测函数 f。

证明 先证明存在子序列 $\{g_k\}$，它几乎处处收敛于 f。

由于 $\{f_n\}$ 是依测度收敛的 Cauchy 列，那么

$$\lim_{m,n\to\infty} \mu\left(|f_n - f_m| \geqslant \frac{1}{2}\right) = 0$$

故存在自然数 n_1, n_2，使得

$$\mu\left(|f_{n_1} - f_{n_2}| \geqslant \frac{1}{2}\right) < \frac{1}{2}$$

取 $g_1 = f_{n_1}, g_2 = f_{n_2}$，即得

$$\mu\left(|g_1 - g_2| \geqslant \frac{1}{2}\right) < \frac{1}{2}$$

一般地说，可取自然数 n_k, n_{k+1}，使

$$\mu\left(|f_{n_k} - f_{n_{k+1}}| \geqslant \frac{1}{2^k}\right) < \frac{1}{2^k}$$

取 $g_k = f_{n_k}$，$g_{k+1} = f_{n_{k+1}}$，令

$$E_k = \left\{\omega \in \Omega: |g_{k+1} - g_k| \geqslant \frac{1}{2^k}\right\}$$

那么

$$\mu(E_k) < 2^{-k}$$

令

$$F_k = \bigcup_{j=k}^{\infty} E_j, \quad F_k \in \mathscr{F}$$

那么

$$\mu(F_k) \leqslant \sum_{j=k}^{\infty} \mu(E_j)$$

$$< \frac{1}{2^k} + \frac{1}{2^{k+1}} + \cdots = 2^{-(k-1)}$$

如果 $i \geqslant j \geqslant k$ 且 $\omega \notin F_k$，那么

$$|g_i - g_j| \leqslant |g_i - g_{i+1}| + \cdots + |g_{j-1} - g_j|$$

$$< \frac{1}{2^{i-1}} + \cdots + \frac{1}{2^j} < \frac{1}{2^{i-2}}$$

又 $\{F_k\}$ 是一递减序列，所以

$$\lim_{k\to\infty} \mu(F_k) = \mu(F) = 0$$

这里 F 是 F_k 的极限。

若 $\omega \in \Omega - F$，则 $\omega \notin F_k$，对于某一 k，因此 g_j 在 $\Omega - F$ 上是收敛序列。定义

$$f = \begin{cases} \lim g_j, & \omega \notin F \\ 0, & \omega \in F \end{cases}$$

这就证明了 g_k 几乎处处收敛于 f。

下面证明 $\{g_j\}$ 依测度收敛于 f。设 α，ε 是两个正实数，选取 k 使得 $\mu(F_k) < 2^{-(k-1)} < \inf(\alpha, \varepsilon)$。若 $j \geqslant k$，那么

$$\left\{\omega \in \Omega: |f(\omega) - g_j(\omega)| \geqslant \alpha\right\} \subset \{\omega \in \Omega: |f(\omega) - g_j|$$

$$> 2^{-(k-1)}\} \subset F_k$$

所以

$$\mu\{\omega\in\Omega: |f(\omega)-g_i(\omega)|\geqslant\alpha\}<\mu(F_k)<\varepsilon \quad \forall j\geqslant k$$

即 g_i 依测度收敛于 f. ▮

推论 1.7.2 设 $\{f_n\}$ 是依测度收敛的 Cauchy 列。那么必存在可测函数 f，使得 f_n 依测度收敛于 f，f 几乎处处唯一确定。

证明 上面我们已经证明了 $\{f_n\}$ 有一子序列 f_{n_k} 依测度收敛于 f. 现在要证明 $\{f_n\}$ 依测度收敛于 f. 因为

$$|f-f_n|\leqslant|f-f_{n_k}|+|f_{n_k}-f_n|$$

所以

$$\{|f-f_n|\geqslant\alpha\}\subset\left\{|f-f_{n_k}|\geqslant\frac{\alpha}{2}\right\}\cup\left\{|f_{n_k}-f_n|\geqslant\frac{\alpha}{2}\right\}$$

由于 $\{f_n\}$ 是 Cauchy 列，f_{n_k} 依测度收敛于 f，因此得

$$\lim_{n\to\infty}\mu\{|f-f_n|\geqslant\alpha\}=0$$

即 f_n 依测度收敛于 f.

假定 $\{f_n\}$ 依测度收敛于 f 和 g，因为

$$|f-g|\leqslant|f-f_n|+|f_n-g|$$

于是有

$$\{|f-g|\geqslant\alpha\}\subset\left\{|f-f_n|\geqslant\frac{\alpha}{2}\right\}\cup\left\{|f_n-g|\geqslant\frac{\alpha}{2}\right\}$$

因此有

$$\mu\{|f-g|\geqslant\alpha\}=0,\ \forall\ \alpha>0$$

取 $\alpha=\dfrac{1}{n}$，$n\in N$，所以 $f=g$ a.e. ▮

我们已经说过，(p 次)平均收敛包含了依测度收敛．一般地说，依测度收敛并不能推出平均收敛(习题 1.7,4)．然而在下面定理的条件下却是成立的．

定理 1.7.4 设 $f_n \in L_p$, $n = 1,2,\cdots$, f_n 依测度收敛于 f，且存在 $g \in L_p$ 使

$$|f_n| \leqslant g, \quad \text{a e.}$$

因此 $f \in L_p$，且 f_n (p 次)平均收敛于 f.

证明 若 f_n 不平均收敛于 $f \in L_p$，那么必有子序列 g_k 和 $\varepsilon > 0$ 使得

$$\|g_k - f\|_p > \varepsilon, \quad \forall k \in N$$

因为 $\{g_k\}$ 是 $\{f_n\}$ 的子序列(习题 1.7,7)，g_k 依测度收敛于 f，由定理 1.7.3 知，存在 $\{g_k\}$ 的子序列 $\{h_r\}$，h_r 几乎处处且依测度收敛于 h.

再由推论 1.7.2 知 $h = f$, a.e.．因为 h_r 几乎处处收敛于 f 且被 g 所控制，由定理 1.7.2，有

$$\|h_r - f\|_p \to 0, \quad 当 \ r \to \infty$$

这与开始得到的不等式矛盾，于是定理得证． ∎

定义 1.7.3 设 $\{f_n\}$ 是一列实可测函数，f 是 Ω 上的实可测函数．如果对于每一个 $\delta > 0$，存在一个可测集 E_δ, $E_\delta \subset \Omega$, $\mu(E_\delta) < \delta$，使 f_n 在 $\Omega - E_\delta$ 上均匀(一致)收敛于 f，则称 $\{f_n\}$ 在 Ω 上几乎均匀(一致)收敛于 f.

从定义可以看出：均匀收敛包含着几乎均匀收敛，反之则不然．

引理 1.7.1 设 $\{f_n\}$ 是几乎均匀的 Cauchy 列，那么必存在可测函数 f 使得 $\{f_n\}$ 几乎均匀收敛且几乎处处收敛于 f.

证明 若 $k \in N$, $E_k \in \mathscr{F}$, $\mu(E_k) < 2^{-k}$，则 $\{f_n\}$ 在 $\Omega - E_k$ 上均匀收敛．若 $F_k = \bigcup_{j=k}^{\infty} E_j$, $F_k \in \mathscr{F}$，且 $\mu(F_k) <$

$2^{-(k-1)}$. 则 $\{f_n\}$ 在 $\Omega - F_k \subset \Omega - E_k$ 上均匀收敛，g_k 定义为

$$g_k(\omega) = \begin{cases} \lim f_n(\omega), & \omega \notin F_k \\ 0, & \omega \in F_k \end{cases}$$

显然 $\{F_k\}$ 是一递减序列，$F = \bigcap F_k$，$\mu(F) = 0$. 若 $h \leqslant k$，那么 $g_h(\omega) = g_k(\omega)$，$\forall \omega \in F_h$.

所以$\{g_k\}$在Ω上收敛于一可测函数，它可记作 f. 若 $\omega \in F_k$，那么 $f = g_k = \lim f_n$. 因为 f_n 在 $\Omega - F$ 上收敛于 f，所以$\{f_n\}$在Ω上几乎处处收敛于 f.

现在证明这一收敛是几乎一致的：对 $\varepsilon > 0$，若K充分大，使得 $2^{-(K-1)} < \varepsilon$. 于是 $\mu(F_K) < \varepsilon$，$\{f_n\}$ 均匀收敛于 $g_k = f$，$\forall \omega \in \Omega - F_K$. ▌

下面定理说明依测度收敛与几乎一致收敛间的关系.

定理1.7.5 若 $\{f_n\}$ 几乎一致收敛于 f，那么 $\{f_n\}$ 依测度收敛于f. 反之，若$\{h_n\}$依测度收敛于 h，那么存在一子序列几乎一致收敛于 h.

证明 $\{f_n\}$ 几乎一致收敛于 f：设 ε，α 是两个任意正实数，那么存在 $E_\varepsilon \in \mathscr{F}$，$\mu(E_\varepsilon) < \varepsilon$ 使得$\{f_n\}$在 $\Omega - E_\varepsilon$ 上一致收敛于 f. 所以对于充分大的 n 有

$$(|f_n - f| \geqslant \alpha) \subset E_\varepsilon$$

或者

$$\mu(|f_n - f| \geqslant \alpha) \leqslant \mu(E_\varepsilon) < \varepsilon$$

即 $\{f_n\}$ 依测度收敛于 f.

反之，若 $\{h_n\}$ 依测度收敛于 h，那么同定理 1.7.3 的证明一样，存在子序列 g_k，g_k 几乎处处且依测度收敛于 f. 用定理 1.7.3 中的记号，若 $j \geqslant k$ 且 $\omega \notin F_k$，那么

$$|h - g_j| \leqslant \frac{1}{2^{j-1}} \leqslant \frac{1}{2^{k-1}}$$

因此 g_k 在 $\Omega - F_k$ 上均匀收敛于 h. ∎

由定理 1.7.5 知，(p 次)平均收敛的序列必存在几乎均匀收敛的子序列。反过来，几乎均匀收敛并不包含平均收敛，如果函数列受到 L_p 中的函数控制，那么几乎均匀收敛是可以推出平均收敛的(利用定理 1.7.4)。

作为引理 1.7.1 的推论，几乎一致收敛包含几乎处处收敛。一般地说，反之不然，(见习题 1.7，12)。然而值得注意的是，如果函数是实值可测函数，$\mu(\Omega) < +\infty$，那么几乎处处收敛包含着几乎一致收敛。这就是著名的 Egoroff 定理。

定理 1.7.6 (Egoroff 定理) 设 $\mu(\Omega) < +\infty$，$\{f_n\}$ 是 Ω 的实值可测函数列，如果 $\{f_n\}$ 在 Ω 上几乎处处收敛于实值可测函数 f，则 $\{f_n\}$ 在 Ω 上几乎一致收敛且依测度收敛于 f.

证明 不失一般性 假设 $\{f_n\}$ 在 Ω 上的每一点都收敛于 f，对于 $m, n \in N$，令

$$E_n(m) = \bigcup_{k=n}^{\infty}\left\{|f_k(\omega) - f(\omega)| \geq \frac{1}{m}\right\}$$

则 $E_n(m) \in \mathscr{F}$ 且 $E_{n+1}(m) \subset E_n(m)$.

因为

$$f_n(\omega) \to f(\omega), \ \forall \omega \in \Omega$$

于是

$$\bigcap_{n=1}^{\infty} E_n(m) = \varnothing$$

又因为 $\mu(\Omega) < +\infty$，所以 $\lim\limits_{n\to\infty}\mu(E_n(m)) = 0$.

往证.若 $\delta > 0$，k_m 使 $\mu(E_{k_m}(m)) < \delta/2m$，而且令

$$E_\delta = \bigcup_{m=1}^{\infty} E_{k_m}(m)$$

那么 $E_\delta \in \mathscr{F}$ 且 $\mu(E_\delta) < \delta$.

显然,若 $\omega \notin E_\delta$, 那么 $\omega \in E'_{k_m}(m)$ 且

$$|f_k(\omega) - f(\omega)| < \frac{1}{m}, \forall k > k_m$$

于是 $\{f_k\}$ 在 E_δ 的补集上均匀收敛. ∎

现在我们用下面的图表来总结各种收敛间的相互关系. 为此,我们引进下列记号:

AE 表示几乎处处收敛;

AU 表示几乎一致收敛;

L_p 表示 p 次平均收敛;

M 表示测度收敛.

图中实线箭头表示包含关系 $A \to B$, 即由 A 可以推出 B. 虚线箭头表示有一个子序列存在此关系.

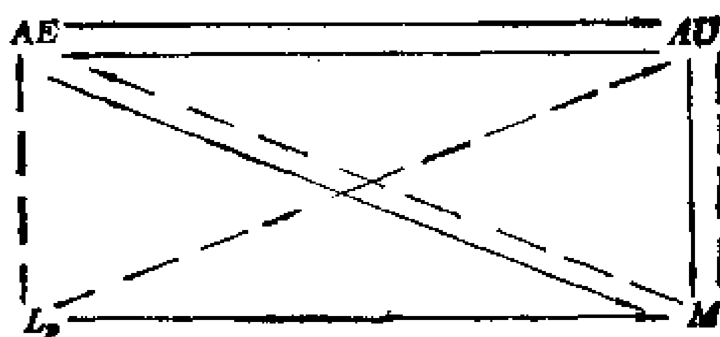


图 1.7.1 一般情况

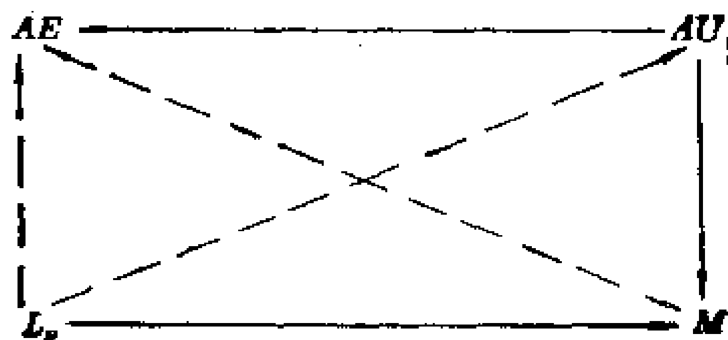


图 1.7.2 有限测度空间

对于有限测度空间 根据 Egoroff 定理又可增加两个包含关系，如图 1.7.2 所示。

如果假设可测函数列 $\{f_n\}$ 被 $g \in L_p$ 所控制，即 $|f_n| \leqslant g$，那么又可增加三种包含关系，见图 1.7.3。

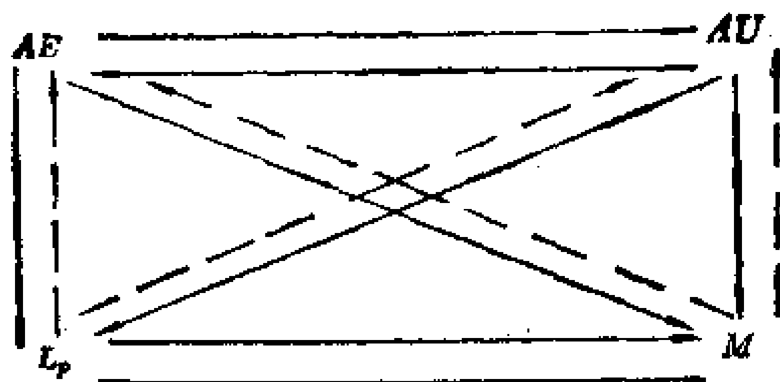


图1.7.3 $\mu(\Omega)<+\infty$，且 $|f_n|\leqslant g$，$g\in L_p$

我们用（p 次）平均收敛的充分必要条件来结束本节。读者将会看到：下列定理中的条件(2)和(3)，在 $|f_n| \leqslant g$，$g \in L_p$ 时，将自动地满足。

定理 1.7.7 (Vitali 定理) 设 $f_n \in L_p$，$1 \leqslant p < \infty$，$n=1,2,\cdots$。f_n 平均（p 次）收敛于 f 的充分必要条件为

(1) $\{f_n\}$ 依测度收敛于 f；

(2) 对每一个 $\varepsilon > 0$，存在集 $\mu(E_\varepsilon) < +\infty$，$E_\varepsilon \in \mathscr{F}$，使得 $F \cap E_\varepsilon = \varnothing$，$F \in \mathscr{F}$，那么

$$\int_F |f_n|^p d\mu < \varepsilon^p,\ \forall\ n \in N$$

(3) 对每一个 $\varepsilon > 0$，存在 $\delta(\varepsilon) > 0$，使得 $\mu(E) < \delta(\varepsilon)$，$E \in \mathscr{F}$，那么

$$\int_E |f_n|^p d\mu < \varepsilon^p,\ \forall\ n \in N$$

证明 在给出了定义 1.7.2 以后，我们就证明了平均（p 次）收敛包含着依测度收敛。因此不难证明$\{f_n\}$平均（p 次）收

敛包含了条件(2)和(3). (习题1.7 的18, 19).

现在,我们证明定理中的三个条件包含着 $\{f_n\}$ 平均收敛于 $f\in L_p$. 若 $\varepsilon>0$, E_ε 是条件(2)给出的集,且 $F=\Omega-E_\varepsilon$.

因为

$$f_n-f_m=(f_n-f_m)\chi_{E_\varepsilon}+(f_n-f_m)\chi_F$$

利用 Minkowshi 不等式,得

$$\|f_n-f_m\|_p\leqslant\left\{\int_{E_\varepsilon}|f_n-f_m|^p d\mu\right\}^{1/p}+2\varepsilon,\ \forall n,\ m\in N$$

令 $\alpha=\varepsilon[\mu(E_\varepsilon)]^{-1/p}$, $H_{n,m}=\{|f_n-f_m|\geqslant\alpha\}$ 根据条件(1),存在 $K(\varepsilon)$,使得在 $n,m\geqslant K(\varepsilon)$ 时

$$\mu(H_{n,m})<\delta(\varepsilon)$$

再次利用 Minkowshi 不等式,并根据条件(3)则在 n, $m\geqslant K(\varepsilon)$ 时有

$$\begin{aligned}\left\{\int_{E_\varepsilon}|f_n-f_m|^p d\mu\right\}^{1/p}&\leqslant\left\{\int_{E_\varepsilon-H_{n,m}}|f_n-f_m|^p d\mu\right\}^{1/p}\\&\quad+\left\{\int_{H_{n,m}}|f_n|^p d\mu\right\}^{1/p}\\&\quad+\left\{\int_{H_{n,m}}|f_m|^p\right\}^{1/p}\\&\leqslant\alpha[\mu(E_\varepsilon)]^{1/p}+\varepsilon+\varepsilon=3\varepsilon\end{aligned}$$

由此可知 $\{f_n\}$ 是平均收敛 Cauchy 列. 已知 $\{f_n\}$ 依测度收敛于 f,由推论 1.7.2 知 f 几乎处处唯一,因此 $\{f_n\}$ 平均收敛于 f. ▌

怎样才能掌握好这些错综复杂的关系呢? 这里有三点可供参考:

(1) 首先应熟记每一种收敛关系的含义, 在你遇到某种收敛关系时, 脑子里就有清晰的概念浮现出来. 这是最基本的一点.

(2) 对于各种收敛关系的强弱程度，心中应该有数，如四种收敛关系中 AU 较强，M 最弱，由 AU 可推出 AE 收敛与 M 收敛，而 L_p 与 AU 均包含 M。

(3) 记住一些著名的定理这可帮助你理清它们之间的关系。Riesz 定理给出 M 收敛列必有 AE 收敛子列；Egoroff 定理给出了 AE 收敛转化为 AU 收敛的条件——$\mu(\Omega)<+\infty$；Vitali 定理给出了平均收敛的充分必要条件。

做到了以上三点，如果又能牢牢掌握积分收敛定理，那么你在处理这类问题时就不会发生谬误。

习　　题　　1.7

在下列习题中 $(R,\mathscr{B},\lambda)$ 表示 Lebesgue 测度空间，即 $R=(-\infty,+\infty)$，$\mathscr{B}$ 是 Lebesgue 可测集类，λ 是 Lebesgue 测度，且假定 $1\leqslant p<\infty$。

1. 令 $f_n=n^{-1/p}\chi_{[0,n]}$，证明 $\{f_n\}$ 均匀收敛于 0 函数，但它不平均收敛(在 L_p 中)。

2. 令 $f_n=n\chi_{[1/n,2/n]}$，证明 $\{f_n\}$ 几乎处处收敛于 0 函数，但它在 L_p 中不平均收敛。

3. 证明习题 1.7，1 和 2 中的函数列依测度收敛于它们的极限。

4. 令 $f_n=\chi_{[n,n+1]}$，证明 $\{f_n\}$ 几乎处处收敛于0函数，但它不依测度收敛。

5. 证明在习题 1.7，2 中的序列测度收敛不包含平均(p 次)收敛，即使 $\mu(\Omega)<+\infty$ 也如此。

6. 找出例 1 中几乎处处收敛于 0 函数的子序列。能否找出一个点点收敛的子序列？

7. 若 $\{f_n\}$ 依测度收敛于 f，试证明 $\{f_n\}$ 的任一子序列也依测度收敛于 f。更一般地，若 $\{f_n\}$ 是依测度收敛的 Cauchy 列，证明 $\{f_n\}$ 的任一子序列也是依测度收敛的 Cauchy 列。

8. 设 $\{f_n\}$ 平均收敛于 f，f_n 的一个子序列平均收敛于 g，试证明 $f=g$，a.e.。

9. 若 $\{f_n\}$ 是 $\mathscr{S}$ 中集的示性函数序列，且 $\{f_n\}$ 平均收敛 (L_p) 于 f，证明 f 是(几乎处处等于) $\mathscr{S}$ 中集的示性函数。

10. 证明习题 1.7,2 中的序列具有下列性质：若 $\delta>0$，f_n 在 $[0,\delta]$ 的补集上均匀收敛；但不存在一个零测集，在它的补集上 $\{f_n\}$ 均匀收敛。

11. 证明习题 1.7,2 中的序列几乎均匀收敛，但不平均收敛。

12. 证明习题 1.7,4 中的序列几乎处处收敛，但不几乎均匀收敛。

13. 证明若用依测度收敛来代替几乎处处收敛，Fatou 引理仍成立。

14. 设 $f_n = n\chi_{(0,\frac{1}{n})}$，证明在 Egoroff 定理中极限函数有限的假设是不能去掉的。

15. 证明若用依测度收敛代替几乎处处收敛，Lebesgue 控制收敛定理仍然成立。

16. 若 $g \in L_p$，$|f_n|\leqslant g$，证明 Vitali 定理中的条件(2)，(3)将满足。

17. 设 $(\Omega,\mathscr{F},\mu)$ 是有限测度空间，f 是可测函数，令

$$r(f)=\int\frac{|f|}{1+|f|}\,d\mu$$

证明当且仅当 $r(f_n-f)\to 0$ 时，$\{f_n\}$ 依测度收敛于 f。

18. 设 $\{f_n\}$ 几乎处处收敛于可测函数 f。φ 是 R 到 R 的连续函数，证明 $\{\varphi\circ f_n\}$ 几乎处处收敛于 $\varphi\circ f$。反之，若 φ 不是每一点都连续，证明存在一个序列 $\{f_n\}$，它几乎处处收敛于 f，但 $\varphi\circ f_n$ 不几乎处处收敛于 $\varphi\circ f$。

19. 若 φ 是 R 到 R 的一致连续函数，且 $\{f_n\}$ 均匀(几乎均匀、依测度)收敛于 f，试证明 $\{\varphi\circ f_n\}$ 均匀(几乎均匀、依测度)收敛于 $\varphi\circ f$。

1.8 测度的分解

提要 本节先引进广义测度的 Hahn 分解和 Jordan 分解，然后定义测度的奇异性。Radon-Nikodym 定理在概率论中具有特殊的重要地位。在

证明了 Radon-Nikodym 定理后，我们证明了：任一σ有限测度λ关于另一σ有限测度μ，可以分解成一个关于μ绝对连续的测度与一个关于μ奇异的测度之和——这就是著名的 Lebesgue 分解定理。最后利用 Lebesgue 分解，讨论任意定分布函数的分解。

广义测度的分解和 Radon-Nikodym 定理是测度论的重要内容之一，在概率论中更具有特殊重要的地位。因此学好这一节，对于本书以后的学习是至关重要的。

在 1.5 节中引进了一般可测函数的积分以后，我们曾经定义过广义测度。广义测度是一个广义实值集函数（它只能是$\pm\infty$之一）。空集的广义测度为零，广义测度具有可列可加性，但不像测度那样只取非负的值，而可以取正值或负值。为了阅读方便，这里我们重述广义测度的定义。

设$(\Omega,\mathscr{F})$为可测空间，λ是定义在$\mathscr{F}$上的广义实值函数(只取$\pm\infty$之一)，且满足

(1) $\lambda(\varnothing)=0$;

(2) 若$\{E_n\}$是$\mathscr{F}$中的不相交集的序列，那么

$$\lambda\left(\bigcup_{n=1}^{\infty}E_n\right)=\sum_{n=1}^{\infty}\lambda(E_n)$$

利用引理 1.3.1 和引理 1.3.2 的方法，读者可以证明：如果$\{E_n\}$是$\mathscr{F}$中的增序列，那么

$$\lambda\left(\bigcup_{n=1}^{\infty}E_n\right)=\lim\lambda(E_n) \tag{1.8.1}$$

如果$\{F_n\}$是$\mathscr{F}$中的减序列，那么

$$\lambda\left(\bigcap_{n=1}^{\infty}F_n\right)=\lim\lambda(F_n) \tag{1.8.2}$$

在本节中，我们首先要证明：任何一个广义测度可以写成两个测度之差。为此我们要引进正集与负集的概念。

定义 1.8.1 设 μ 是 $(\Omega, \mathscr{F})$ 上的广义测度. $A \in \mathscr{F}$，若对任何 $E \in \mathscr{F}$ 有 $\mu(A \cap E) \geqslant 0$，则称 A 为正集（关于 μ 而言）。若 $B \in \mathscr{F}$，对任何 $E \in \mathscr{F}$ 有 $\mu(B \cap E) \leqslant 0$，则称 B 为负集。若 $C \in \mathscr{F}$，对任何 $E \in \mathscr{F}$ 有 $\mu(C \cap E) = 0$，则称 C 为零集。

在习题中我们将看到，正集的可测子集是正集，两个正集之和是正集。

定理 1.8.1（Hahn 分解定理） 若 μ 是 $\mathscr{F}$ 上的广义测度，则存在两个不相交的集合 A 和 B，使得

$$A \cup B = \Omega$$

且 A 对于 μ 为正集，B 对于 μ 为负集。

证明 因为 μ 在 $+\infty$ 与 $-\infty$ 两个值中只能取一个值。为确定起见，我们假定 μ 不取 $-\infty$，即对于任何 $E \in \mathscr{F}$

$$-\infty < \mu(E) \leqslant +\infty$$

令

$$\beta = \inf\{\mu(B) | B \text{ 是 } \mu \text{ 负集}\}$$

则存在负集序列 $B_n \in \mathscr{F}$，$n = 1, 2, \cdots$，使得

$$\lim_{n \to \infty} \mu(B_n) = \beta$$

令 $B = \bigcup_{n=1}^{\infty} B_n$，则 B 是一个负集，因此 $\beta \leqslant \mu(B)$。另一方面，因为 B 是负集，故 $\mu(B - B_n) \leqslant 0$，所以

$$\mu(B) = \mu(B_n) + \mu(B - B_n) \leqslant \mu(B_n)$$

令 $n \to \infty$ 可得

$$\mu(B) \leqslant \beta$$

从而得到 $\mu(B) = \beta$（即负集的 μ 值可以达到最小）。

现在我们证明$A=\Omega-B$是一个正集。假设结论不成立,设 E_0 是A的一个可测子集且使得 $\mu(E_0)<0$。我们首先看出 E_0 不可能是负集。因为如果 E_0 是负集,那么 $B\cup E_0$ 也将是负集,且

$$\mu(B\cup E_0)=\mu(B)+\mu(E_0)<\mu(B)=\beta$$

这与负集的μ值可以达到最小相矛盾。因此 E_0 存在μ值为正的可测子集,令

$$k_1=\min\left\{n\,\middle|\,\text{存在可测集 } E\subset E_0,\ \text{使得 } \mu(E)\geqslant\frac{1}{n}\right\}$$

并设 E_1 是 E_0 的可测子集,使得 $\mu(E_1)\geqslant\frac{1}{k_1}$。因为 $\mu(E_0)<0$,所以

$$\begin{aligned}\mu(E_0-E_1)&=\mu(E_0)-\mu(E_1)\\&\leqslant\mu(E_0)-\frac{1}{k_1}<0\end{aligned}$$

刚才用于 E_0 的论证也可以同样用于 E_0-E_1,即令

$$k_2=\min\left\{n\,\middle|\,\text{存在可测集 } E\subset E_0-E_1,\text{使得 } \mu(E)\geqslant\frac{1}{n}\right\}$$

并设 E_2 是 E_0-E_1 的可测子集,使得 $\mu(E_2)\geqslant\frac{1}{k_2}$,则

$$\begin{aligned}\mu(E_0-E_1-E_2)&=\mu(E_0)-\mu(E_1)-\mu(E_2)\\&\leqslant\mu(E_0)-\frac{1}{k_1}-\frac{1}{k_2}<0\end{aligned}$$

依此类推,令

$$F_0=E_0-\bigcup_{n=1}^{\infty}E_n$$

由于 $\mu(F_0)$ 有限且

$$\mu(F_0)=\mu(E_0)-\sum_{n=1}^{\infty}\mu(E_n)$$

$$\leqslant \mu(E_0) - \sum_{n=1}^{\infty} \frac{1}{k_n}$$

故 $\sum_{n=1}^{\infty} \frac{1}{k_n}$ 收敛，因而有

$$\lim_{n \to \infty} \frac{1}{k_n} = 0$$

由此可知 F_0 不可能存在 μ 值为正的可测子集，即如果 F 是 F_0 的任意子集，则有 $\mu(F) \leqslant 0$。这也就是说，F_0 是一个负集，因而 $B \cup F_0$ 也是负集。但 F_0 与 B 不相交，因为

$$\mu(F_0) \leqslant \mu(E_0) < 0$$

于是

$$\mu(B \cup F_0) = \mu(B) + \mu(F_0) < \mu(B) = \beta$$

这与 $\mu(B)$ 的最小性质矛盾，故 A 是正集且

$$A \cup B = \Omega$$ ▌

设 A 与 B 是一对可测集，满足上面定理的结论，于是 A，B 称为 Ω 关于 μ 的一个 Hahn 分解。

一般地说，Hahn 分解不是唯一的。事实上，如果 A，B 是关于 μ 的 Hahn 分解，C 是 μ 零集，那么 $A \cup C$，$B - C$ 和 $A - C$，$B \cup C$ 也都是 Hahn 分解。然而，在下面的讨论中 Hahn 分解没有唯一性并不重要。

引理 1.8.1 若 A_1，B_1 和 A_2，B_2 都是关于 λ 的 Hahn 分解，且 $E \in \mathscr{F}$，那么

$$\lambda(E \cap A_1) = \lambda(E \cap A_2)$$
$$\lambda(E \cap B_1) = \lambda(E \cap B_2) \tag{1.8.3}$$

证明 因为 $E \cap (A_1 - A_2)$ 包含于正集 A_1 和负集 B_2 中，所以

$$\lambda(E \cap (A_1 - A_2)) = 0$$
$$\lambda(E \cap A_1) = \lambda(E \cap A_1 \cap A_2) + \lambda(E \cap (A_1 - A_2))$$

$$= \lambda(E \cap A_1 \cap A_2)$$

同理

$$\lambda(E \cap A_2) = \lambda(E \cap A_1 \cap A_2)$$

因此

$$\lambda(E \cap A_1) = \lambda(E \cap A_2)$$

式(1.8.3)同样证明. ∎

定义 1.8.2 设 λ 是 $\mathscr{F}$ 上的实值广义测度（不取无穷值）. A, B 是 λ 的 Hahn 分解，对于 $E \in \mathscr{F}$，有

$$\lambda^+(E) = \lambda(E \cap A), \quad \lambda^-(E) = -\lambda(E \cap B) \quad (1.8.4)$$

λ^+ 和 λ^- 都是 $\mathscr{F}$ 上的有限测度，称为 λ 的正变差，负变差.

$$|\lambda|(E) = \lambda^+(E) + \lambda^-(E)$$

称为 λ 的全变差，记作 $|\lambda|$.

由引理 1.8.1 可知，正变差和负变差是完全确定的，不依赖于不同的 Hahn 分解. 同时，根据式(1.8.4)有

$$\lambda(E) = \lambda(E \cap A) + \lambda(E \cap B)$$
$$= \lambda^+(E) - \lambda^-(E) \quad (1.8.5)$$

由此可知，任何实值广义测度可以写成两个测度之差.

定理 1.8.2 (Jordan 分解定理) λ 是 $\mathscr{F}$ 上的广义测度，也是 $\mathscr{F}$ 上的两个有限测度之差，即 $\lambda = \lambda^+ - \lambda^-$. 如果 $\lambda = \mu - \nu$，其中 μ, ν 是 $\mathscr{F}$ 上的有限测度，那么

$$\mu(E) \geqslant \lambda^+(E), \quad \nu(E) \geqslant \lambda^-(E), \quad \forall E \in \mathscr{F} \quad (1.8.6)$$

证明 由条件 $\lambda = \lambda^+ - \lambda^-$ 的表示式看出，已经成立. 因为 μ, ν 只取非负的值，因此

$$\lambda^+(E) = \lambda(E \cap A) = \mu(E \cap A) - \nu(E \cap A)$$
$$\leqslant \mu(E \cap A) \leqslant \mu(E)$$

类似地可证明 $\lambda^-(E) \leqslant \nu(E)$，$\forall E \in \mathscr{F}$. ∎

在引理 1.5.1 中，我们已经看到：如果 f 是关于 μ 的可积函数，且

$$\lambda(E)=\int_E f d\mu,\ \forall E\in\mathscr{F} \tag{1.8.7}$$

那么 λ 是 $\mathscr{F}$ 上的广义测度。

现在我们进一步来讨论正变差，负变差与积分的关系。

定理 1.8.3 若 $f\in L(\Omega,\mathscr{F},\mu)$，$\lambda$ 是由式(1.8.7)定义的广义测度，那么正变差 λ^+，负变差 λ^- 和全变差 $|\lambda|$ 给出如下：如果 $\forall E\in\mathscr{F}$，于是

$$\lambda^+(E)=\int_E f^+ d\mu,\qquad \lambda^-(E)=\int_E f^- d\mu$$

$$|\lambda|(E)=\int_E |f| d\mu$$

证明 设 $A_f=\{\omega\in\Omega\,|\,f(\omega)\geqslant 0\}$，$B_f=\{\omega\in\Omega\,|\,f(\omega)<0\}$，那么 $\Omega=A_f\cup B_f$ 且 $A_f\cap B_f=\emptyset$。若 $E\in\mathscr{F}$，显然有 $\lambda(E\cap A_f)\geqslant 0$，$\lambda(E\cap B_f)\leqslant 0$。因此 A_f 和 B_f 就是 λ 的 Hahn 分解。 ∎

由推论 1.4.2 可知，若 f 是非负广义实值可测函数，μ 是 $\mathscr{F}$ 上的测度，那么由式 (1.8.7)定义的函数 λ 是 $\mathscr{F}$ 上的测度。

定义 1.8.3 λ 和 μ 是 $\mathscr{F}$ 上的两个测度，若对于 $E\in\mathscr{F}$，$\mu(E)=0$ 时有 $\lambda(E)=0$，那么 λ 关于 μ 绝对连续。广义测度 λ 和 μ，若对于 $|\mu(E)|=0$ 的每一 $E\in\mathscr{F}$，有 $\lambda(E)=0$，则 λ 关于 μ 绝对连续，记作 $\lambda\ll\mu$。

下面的引理可以帮助理解绝对连续的内含。

引理 1.8.2(A) λ 和 μ 是 $\mathscr{F}$ 上的有限测度。$\lambda\ll\mu$ 的充分必要条件是：对于每一个 $\varepsilon>0$，存在 $\delta(\varepsilon)>0$ 使得 $\mu(E)<\delta(\varepsilon)$，$E\in\mathscr{F}$ 时有 $\lambda(E)<\varepsilon$。

证明 充分性：若条件满足 $\mu(E)=0$，那么 $\lambda(E)<\varepsilon$，对任何 $\varepsilon>0$ 成立，因此 $\lambda(E)=0$。

必要性：用反证法。假设存在 $\varepsilon>0$，$E_n\in\mathscr{F}$ 有

$\mu(E_n) < 2^{-n}$ 且 $\lambda(E_n) \geqslant \varepsilon$，并令 $F_n = \bigcup\limits_{k=n}^{\infty} E_k$，$\mu(F_n) < 2^{-n+1}$ 且 $\lambda(F_n) \geqslant \varepsilon$. 则 F_n 是可测集的递减序列：

$$\mu\left(\bigcap_{n=1}^{\infty} F_n\right) = \lim \mu(F_n) = 0$$

$$\lambda\left(\bigcap_{n=1}^{\infty} F_n\right) = \lim \lambda(F_n) \geqslant \varepsilon$$

这与 λ 关于 μ 绝对连续矛盾，必要性得证。 ▮

引理 1.8.2(B) 设 μ 和 λ 是 $\mathscr{F}$ 上的两个有限测度，$\lambda \ll \mu$ 且 λ 不恒为零，则存在正数 ε 和可测集 A，使得 $\mu(A) > 0$，A 为广义测度 $\lambda - \varepsilon\mu$ 的正集。

证明 设 $\Omega = A_n + B_n$ $(n = 1, 2, \cdots)$ 是 Ω 对广义测度 $\lambda - \dfrac{1}{n}\mu$ 的 Hahn 分解，并令

$$A_0 = \bigcup_{n=1}^{\infty} A_n, \quad B_0 = A_0^c = \bigcap_{n=1}^{\infty} B_n$$

由于 $B_0 \subset B_n$，故有 $\left(\lambda - \dfrac{1}{n}\mu\right)(B_0) \leqslant 0$，即

$$0 \leqslant \lambda(B_0) \leqslant \frac{1}{n}\mu(B_0), n = 1, 2, \cdots$$

从而有 $\lambda(B_0) = 0$. 由于 λ 不恒为零，故 $\lambda(A_0) > 0$，根据 $\lambda \ll \mu$ 可得 $\mu(A_0) > 0$. 因此至少对于 n 的一个值有 $\mu(A_n) > 0$；对于这个 n 值，我们令 $A = A_n$，$\varepsilon = \dfrac{1}{n}$，于是 A，ε 为所求的且 A 为 $\lambda - \varepsilon\mu$ 的正集。 ▮

定理 1.8.4(Radon-Nikodym 定理) 设 λ 和 μ 都是 $\mathscr{F}$ 上的 σ 有限测度. 假定 λ 关于 μ 绝对连续，那么存在一非负可测函数 $f \in M^{+}(\Omega, \mathscr{F})$ 使得

$$\lambda(E)=\int_E f d\mu,\ \forall E\in\mathscr{F} \tag{1.8.7}$$

且 f 几乎处处唯一确定。

证明 分两步进行。我们先在 $\lambda(\Omega)<+\infty$ 的假设下证明定理。

令

$$\kappa=\left\{f\in M^+:\ \int_E f d\mu\leqslant\lambda(E),\ \forall E\in\mathscr{F}\right\}$$

这个函数类是非空的，因为至少 $0\in\kappa$。

(1) 可以证明 κ 关于"取极大"运算封闭。事实上，若 $f,g\in\kappa$，则

$$f\vee g\triangleq\max(f,g)$$

令

$$E_1=\{\omega\in E: f(\omega)\geqslant g(\omega)\}$$

于是

$$\begin{aligned}\int_E(f\vee g)d\mu&=\int_{E_1\cup(E-E_1)}(f\vee g)d\mu\\&=\int_{E_1}fd\mu+\int_{E-E_1}gd\mu\\&\leqslant\lambda(E_1)+\lambda(E-E_1)=\lambda(E)\end{aligned}$$

所以 $f\vee g\in\kappa$。

(2) 因为 κ 非空，则设存在上确界

$$\sup_{f\in\kappa}\int fd\mu=\alpha$$

α 是有限数。因此存在一个序列 $\{f_n\}\subset\kappa$，使得

$$\lim_n\int f_n d\mu=\alpha$$

作函数 g_n，即

$$g_n=f_1\vee f_2\vee\cdots\vee f_n\in\kappa$$

则 $\{g_n\}$ 是上升序列，$\lim g_n=f_0$ 且

$$\int g_n d\mu \leqslant \alpha$$

根据定理 1.4.1（单调收敛定理），可知

$$\lim \int g_n d\mu = \int f_0 d\mu \leqslant \alpha$$

所以 $f_0 \in \kappa$。

（3）证明 $\int_E f_0 d\mu = \lambda(E)$，采用反证法。

若 $\lambda(E) - \int_E f_0 d\mu \neq 0$，如果 $\lambda(E) - \int_E f_0 d\mu > 0$，那么根据引理 1.8.2(B)，存在 ε 和 A，A 为 $\left[\lambda(\cdot) - \int_{(\cdot)} f_0 d\mu - \varepsilon\mu\right]$ 正集，即

$$\lambda(E) - \int_E f_0 d\mu - \varepsilon \int_E \chi_A d\mu \geqslant 0, \forall E \in \mathscr{F}$$

$$\lambda(E) \geqslant \int_E (f_0 + \varepsilon \chi_A) d\mu$$

故 $f_0 + \varepsilon\chi_A \in \kappa$，其中 $\mu(A) > 0$。

现在计算

$$\int (f_0 + \varepsilon\chi_A) d\mu = \int f_0 d\mu + \varepsilon\mu(A)$$

$$= \alpha + \varepsilon\mu(A) > \alpha$$

这与 α 是 κ 中函数积分的上确界矛盾，故

$$\lambda(E) = \int_E f_0 d\mu$$

往证。μ，λ 是 σ 有限测度时 Radon-Nikodym 定理仍然成立。

设 $\{\Omega_n\}$ 是 $\mathscr{F}$ 中的增序列，使得

$$\lambda(\Omega_n) < +\infty, \ \mu(\Omega_n) < +\infty$$

利用上面的讨论，取 $g_n \in M^+$，g_n 在 $\omega \notin \Omega_n$ 时等于零，于是 E 是 Ω_n 的子集，且

$$\lambda(E)=\int_E g_n d\mu$$

若 $n\leqslant m$，$\Omega_n\subset\Omega_m$ 有

$$\int_E g_n d\mu=\int_E g_m d\mu, \forall E\subset\Omega_n, \ E\in\mathscr{F}$$

则 g_m 是几乎处处唯一的（这一点最后证），所以当 $m\geqslant n$ 时 $g_m=g_n$，a.e.。

令 $f_n=\sup\{g_1, \cdots, g_n\}$，$f_n$ 是单调增序列，$f_n\in M^+$，$\lim f_n=f$. 若 $E\in\mathscr{F}$，那么

$$\lambda(E\cap\Omega_n)=\int_E f_n d\mu$$

因为 $\{E\cap\Omega_n\}$ 是单调增序列，根据测度连续性引理 1.3.2 和单调收敛定理 1.4.1，可知

$$\begin{aligned}\lambda(E)=\lim\lambda(E\cap\Omega_n)&=\lim\int_E f_n d\mu\\&=\int_E f d\mu\end{aligned}$$

现在证明 f 的唯一性：假设 $f, h\in M^+$，使得

$$\lambda(E)=\int_E f d\mu=\int_E h d\mu, \ \forall E\in\mathscr{F}$$

令 $E_1=\{\omega:f(\omega)>h(\omega)\}$，$E_2=\{\omega:f(\omega)<h(\omega)\}$，则

$$\int_E(f-h)d\mu=\int_{E_1\cup E_2}(f-h)d\mu$$

利用推论 1.4.3 可证得 $f=h$，a.e.。 ∎

Radon-Nikodym 定理在概率论中是十分重要的，它保证了条件数学期望的存在性。

从这里我们看到，Radon-Nikodym 定理还保证了存在 $f\in M^+$ 使得

$$\lambda(E)=\int_E f d\mu$$

f 称为 Radon-Nikodym 导数,记作 $d\lambda/d\mu$, 即

$$f=\frac{d\lambda}{d\mu} \tag{1.8.8}$$

在习题中，我们将讨论导数的有关性质。读者应该注意到：一般地说，f 不一定可积,只有当 λ 是有限测度时 f 才可积。

从直观上看，$\lambda\ll\mu$ 表示 μ 测度小的集，其 λ 测度也很小。与这一情况相反的极端情况是奇异测度。下面讨论这个问题。

定义 1.8.4 若对 $\mathscr{F}$ 上的两个测度 λ,μ,存在不相交集 A, B, $A\cup B=\Omega$ 且 $\mu(A)=\lambda(B)=0$ 则称 λ 与 μ 相互奇异,记作 $\lambda\perp\mu$。

显然这种关系对于 λ 和 μ 是对称的，然而有时我们仍称 λ 关于 μ 奇异.

定理 1.8.5 (Lebesgue 分解定理) 若 λ 和 μ 都是 $\mathscr{F}$ 上的 σ 有限测度，那么必存在测度 λ_1 和 λ_2(其中 λ_1 关于 μ 奇异,λ_2 关于 μ 绝对连续),使得 $\lambda=\lambda_1+\lambda_2$, 且 λ_1, λ_2 唯一确定。

证明 令 $\nu=\lambda+\mu$, 因为 λ, μ 是 σ 有限测度,故 ν 也是 σ 有限测度。因为 $\lambda\ll\nu$, $\mu\ll\nu$, 由 Radon-Nikodym 定理可知,存在 $f,g\in M^+(\Omega,\mathscr{F})$ 使得

$$\lambda(E)=\int_E f d\nu,\quad \mu(E)=\int_E g d\nu,\quad \forall E\in\mathscr{F}$$

令 $A=\{\omega:\ g(\omega)=0\}$, $B=\{\omega:\ g(\omega)>0\}$, 因此 $A\cap B=\varnothing$, $A\cup B=\Omega$。

对于 $\mathscr{F}$ 中的集 E,定义

$$\lambda_1(E)=\lambda(E\cap A),\quad \lambda_2(E)=\lambda(E\cap B)$$

因为 $\mu(A)=0$, 所以 $\lambda_1(B)=\lambda(B\cap A)=\lambda(\varnothing)=0$,故

$\lambda_1 \perp \mu$。为了证明 $\lambda_2 \ll \mu$。考虑下列事实：若 $\mu(E) = 0$，那么

$$\int_E g d\nu = 0, \forall E \in \mathscr{F}$$

所以 $g(\omega) = 0$，a.e. $\nu(E \cap B) = 0$。因为 $\lambda \ll \nu$，且

$$\nu(E \cap B) = \lambda(E \cap B) + \mu(E \cap B)$$

故

$$\lambda_2(E) = \lambda(E \cap B) = 0$$

显然 $\lambda = \lambda_1 + \lambda_2$，这就证明了分解的存在性。

证明分解的唯一性：

$$\lambda = \lambda_1 + \lambda_2 = \tilde{\lambda}_1 + \tilde{\lambda}_2$$

此处 $\lambda_1 \perp \mu$，$\tilde{\lambda}_1 \perp \mu$ 且 $\lambda_2 \ll \mu$，$\tilde{\lambda}_2 \ll \mu$。因此存在可测集 $A, B, \tilde{A}, \tilde{B}$ 使得

$$\Omega = A \cup B = \tilde{A} \cup \tilde{B},\quad A \cap B = \varnothing,\quad \tilde{A} \cap \tilde{B} = \varnothing$$

和

$$\lambda_1(B) = \mu(A) = \tilde{\lambda}_1(\tilde{B}) = \mu(\tilde{A}) = 0$$

设 $E \in \mathscr{F}$，则

$$E = (E \cap B \cap \tilde{B}) + (E \cap B \cap \tilde{A}) + (E \cap A \cap \tilde{B}) + (E \cap A \cap \tilde{A})^{1)}$$

显然 μ 在这个和的最后三个集上为 0，故由绝对连续性知，这三个集的 λ_2 与 $\tilde{\lambda}_2$ 的测度也为 0。又因为 $\tilde{\lambda}_2 - \lambda_2 = \lambda_1 - \tilde{\lambda}_1$，$\lambda_1(B) = \tilde{\lambda}_1(\tilde{B}) = 0$，故有

$$(\tilde{\lambda}_2 - \lambda_2)(E) = (\tilde{\lambda}_2 - \lambda_2)(E \cap B \cap \tilde{B}) = (\lambda_1 - \tilde{\lambda})(E \cap B \cap \tilde{B}) = 0$$

从而有 $\lambda_2(E) = \tilde{\lambda}_2(E)$，而且因为 $\lambda_1 - \tilde{\lambda}_1 = \tilde{\lambda}_2 - \lambda_2$ 所以 $\lambda_1(E) = \tilde{\lambda}_1(E)$，于是我们证明了分解定理。 ∎

---

1) 这里用"+"号代替集合运算联，表示不相交集之联。

利用 Lebesgue 分解定理可以讨论分布函数的分解问题。分布函数可定义如下:

定义 1.8.5 设 $F(x)$ 是定义在 $R=(-\infty,\infty)$ 上的实值函数，如果满足条件:

(1) $F(x)$ 单调增加，即当 $x_1<x_2$ 时，有

$$F(x_1)\leqslant F(x_2)$$

(2) $F(x)$ 右连续，即当 $x\in R$ 时，有

$$F(x+0)=\lim_{x'\to x+0}F(x')=F(x)$$

则称 $F(x)$ 是分布函数。

(3) 又若

$$F(-\infty)=\lim_{x\to-\infty}F(x)=0$$

$$F(+\infty)=\lim_{x\to+\infty}F(x)<\infty \tag{1.8.9}$$

则称 $F(x)$ 是定分布函数。

满足条件 $F(+\infty)=1$ 的定分布函数称为概率分布函数。

设 $R=(-\infty,+\infty)$，$\mathscr{B}$ 是 R 中 Borel 集的全体，可以证明对于任意给定的分布函数 $F(x)$，在 $\mathscr{B}$ 上存在唯一的测度 p，使得

$$p((a,b])=F(b)-F(a)$$

p 称为 Lebesgue-Stieltjes 测度。

为了讨论定分布函数的分解，我们还得定义绝对连续函数。

定义 1.8.6 设 $F(x)$ 是定义在实数空间 R 上的实值函数，若对任给的 $\varepsilon>0$，恒存在 $\delta>0$，使得当 R 中任何有限个两两不相交的开区间 $(a_1,b_1),(a_2,b_2),\cdots,(a_n,b_n)$ 满足条件

$$\sum_{k=1}^{n}(b_k-a_k)<\delta$$

时，不等式

$$\sum_{k=1}^{n}|F(b_k)-F(a_k)|<\varepsilon$$

恒成立，则称 F 为 R 上的绝对连续函数。

显然，绝对连续函数是在通常意义下的连续函数（取 $n=1$），但其逆不真，详见下例。

例

令

$$F(x)=\begin{cases}x\cos\dfrac{\pi}{2x}, & x\neq 0\\ 0, & x=0\end{cases}$$

显然 $F(x)$ 是一连续函数，下面我们要证明 $F(x)$ 在区间[0，1]内并不绝对连续。

设 n 与 m 是任意正整数，令

$$a_k=\frac{1}{2(m+k)},\quad b_k=\frac{1}{2(m+k)-1},\quad k=1,2,\cdots,n$$

则

$$\sum_{k=1}^{n}(b_k-a_k)<\frac{1}{2m+1} \tag{1.8.10}$$

$$\sum_{k=1}^{n}|F(b_k)-F(a_k)|\geqslant\sum_{k=1}^{n}\frac{1}{2(m+k)} \tag{1.8.11}$$

由于 m 可以任意大，故由式(1.8.10)知 $\sum\limits_{k=1}^{n}(b_k-a_k)$ 可以任意小；由于 n 可以任意大，故由式（1.8.11）知 $\sum\limits_{k=1}^{n}|F(b_k)-F(a_k)|$ 可以任意大。由此可见，$F(x)$ 不绝对连续。

为了确定起见，当我们谈到 $\mathscr{B}$ 上的有限测度 ν 的分布函数时，就是指如下的定分布函数：

$$F(x)=\nu((-\infty,x])$$

引理1.8.3 设 ν 是 $(R,\mathscr{B})$ 上的有限测度，那么 $\nu\ll\mu$ 的充分必要条件为 ν 的分布函数 $F(x)$ 是 R 上的绝对连续函数。

证明 充分性：设 $F(x)$ 绝对连续，则对任意给定的正数 ε，存在正数 δ，使得当 R 中任何有限个两两不相交的开区间 $(a_k,b_k),(k=1,2,\cdots,n)$ 满足条件 $\sum_{k=1}^{n}(b_k-a_k)<\delta$ 时，有

$$\sum_{k=1}^{n}|F(b_k)-F(a_k)|<\varepsilon$$

设 $E\in\mathscr{B}$ 且 $\mu(E)=0$，由 Lebesgue 测度性质易知，存在两两不相交的半区间 $(a_k,b_k],k=1,2,\cdots$，使得 $E\subset\bigcup_{k=1}^{\infty}(a_k,b_k]$ 且

$$\sum_{k=1}^{\infty}(b_k-a_k)<\delta$$

由此可见，对任意正整数 n，恒有

$$\sum_{k=1}^{n}|F(b_k)-F(a_k)|<\varepsilon$$

因而

$$\nu(E)\leqslant\sum_{k=1}^{\infty}\nu((a_k,b_k])=\sum_{k=1}^{\infty}|F(b_k)-F(a_k)|\leqslant\varepsilon$$

由于 ε 任意，故 $\nu(E)=0$，这就证明了 $\nu\ll\mu$。

必要性：设 $\nu\ll\mu$，则对于每一个正数 ε，存在正数 δ，使得对于满足条件 $\mu(E)<\delta$ 的任何 Borel 集 E，有 $\nu(E)<\varepsilon$。由此可知，当 R 中任何有限个两两不相交的开区间 $(a_k,b_k),k=1,2,\cdots,n$，满足条件

$$\mu\left(\bigcup_{k=1}^{n}(a_k,b_k]\right)=\sum_{k=1}^{n}(b_k-a_k)<\delta$$

时，有

$$\sum_{k=1}^{n} |F(b_k) - F(a_k)| = \sum_{k=1}^{n} \nu((a_k, b_k])$$

$$= \nu\left(\bigcup_{k=1}^{n} (a_k, b_k]\right) < \varepsilon$$

故 $F(x)$ 绝对连续. ∎

引理 1.8.4 设 $F(x)$ 是绝对连续的定分布函数，那么存在非负实可积函数 f，使得

$$F(x) = \int_{-\infty}^{x} f d\mu$$

这里 f 称为 F 的密度函数，μ 是 Lebesgue 测度。

证明 设 μ_F 是 $F(x)$ 引出的 Lebesgue-Stieltjes 测度，于是 $F(x)$ 是 μ_F 的分布函数：

$$\mu_F((-\infty, x]) = F(x)$$

由引理 1.8.3 知，$\mu_F \ll \mu$。于是由 Radon-Nikodym 定理知，存在非负实可积函数 f，使得 $E \in \mathscr{F}$，有

$$\mu_F(E) = \int_E f d\mu$$

令 $E = (-\infty, x]$ 得

$$\mu_F((-\infty, x]) = \int_{-\infty}^{x} f d\mu = F(x)$$ ∎

定义 1.8.7 设 $F(x)$ 是定分布函数，如果存在非负可积函数 f，使得

$$F(x) = \int_{-\infty}^{x} f d\mu$$

则称 $F(x)$ 是连续型分布函数，f 称为 F 的密度。

引理 1.8.5 连续型分布函数一定绝对连续。

证明 设 f 是 F 的密度，令

$$\nu(E)=\int_E f d\mu$$

则 ν 是 $\mathscr{B}$ 上的测度，且 $\nu\ll\mu$。由定义 1.8.7 知 $F(x)$ 是 ν 的分布函数，由引理 1.8.3 知，$F(x)$ 是绝对连续函数。 ▌

由以上两个引理可知，就定分布函数而言，绝对连续的概念与连续型的概念是一致的。 由于连续函数未必绝对连续，所以连续的分布函数未必是连续型分布函数。

定义 1.8.8 设 $F(x)$ 是定分布函数，如果存在 R 中的点列 $\{x_n\}$ 以及定义在其上的实值函数 $p(x_n)$，使得

$$F(x)=\sum_{x_n\leqslant x}p(x_n),\quad p(x_n)\geqslant 0$$

则称 $F(x)$ 是离散型分布函数。

引理 1.8.6 任何定分布函数 F 都可以唯一地分解为

$$F=F_c+F_d$$

其中 F_d 为离散型分布函数，F_c 是连续（不一定是连续型）的定分布函数。

证明 因为 F 是增函数，故它至多有可列个不连续点，设为 x_n，$n=1,2,\cdots$。令

$$p(x_n)=F(x_n)-F(x_n-0)>0$$

$$F_d(x)=\sum_{x_n\leqslant x}p(x_n),\quad x\in R$$

则 $F_d(x)$ 是离散型分布函数。令

$$F_c=F-F_d,\ \forall x\in R$$

显然 $F=F_c+F_d$，且 F_c 是右连续函数。 下面证明 F_c 也是单调增的左连续函数。对于任意的 $x'<x$，有

$$\begin{aligned}F_c(x)-F_c(x')&=F(x)-F(x')-\sum_{x'<x_n\leqslant x}p(x_n)\\&=F(x-0)-F(x')-\sum_{x'<x_n<x}p(x_n)\end{aligned}$$

由此得

$$\lim_{x'\to x-0}[F_c(x)-F_c(x')]=F(x-0)-F(x-0)=0$$

即 F_c 是左连续的.

又根据 $p(x_n)$ 的定义,对任意的 $x'<x$, 有

$$F_c(x)-F_c(x')=F(x)-F(x')-\sum_{x'<x_n<x}p(x_n)$$

$$=[F(x)-F(x')]-\sum_{x'<x_n<x}$$

$$\times[F(x_n)-F(x_n-0)]\geqslant 0$$

所以 F_c 是增函数，从而证明了 F_c 是连续的分布函数. 显然 F_c 也是定分布函数.

证明分解的唯一性：设 F 有两个分解式

$$F=F_c+F_d=\tilde{F}_c+\tilde{F}_d \qquad (1.8.12)$$

其中 F_c 与 $\tilde{F}_c$ 是连续的定分布函数，而 F_d 与 $\tilde{F}_d$ 是离散型分布函数。由式(1.8.12)我们有

$$F_c-\tilde{F}_c=\tilde{F}_d-F_d$$

上式左边为连续函数，右边为两个离散分布函数之差. 容易证明这个差必须为零,即我们有 $\tilde{F}_d=F_d$, $\tilde{F}_c=F_c$. ∎

由这个引理知,任何定分布函数可以唯一地分解为 $F=F_c+F_d$. 用 ν_c 表示由 F_c 引出的 Lebesgue-Stieltjes 测度. 由 Lebesgue 分解定理知 ν_c 对于 Lebesgue 测度 μ 可唯一地分解为

$$\nu_c=\nu_{ac}+\nu_s$$

其中 $\nu_{ac}\ll\mu$, $\nu_s\perp\mu$.

由 Radon-Nikodym 定理知,存在非负实可积函数 f, 使得对每个 $E\in\mathscr{B}$, 有

$$\nu_{ac}(E)=\int_E f\,d\mu$$

令

$$F_{ac}(x)=\nu_{ac}((-\infty,x]),\quad \forall x\in R$$
$$F_s(x)=\nu_s((-\infty,x]),\quad \forall x\in R$$

则 F_{ac} 与 F_s 分别为 ν_{ac} 与 ν_s 的定分布函数。由引理 1.8.3 知 $F_{ac}(x)$ 是连续型分布函数，$F_s(x)$ 为奇异型分布函数。人们已经证明 $F_s(x)$ 是一个导数几乎处处等于零的函数。

综上所述我们得到：

定理 1.8.6（定分布函数分解定理） 任何定分布函数 F 都可唯一地分解为

$$F=F_d+F_{ac}+F_s$$

其中 F_d 为离散型分布函数，F_{ac} 为连续型分布函数，F_s 为奇异型分布函数。

习 题 1.8

1. 若 A 是关于广义测度 λ 的一个正集，$E\in\mathscr{F}$，$E\subset A$，试证明 E 关于 λ 也是正集。

2. 若 A_1 和 A_2 是广义测度 λ 的两个正集，证明 $A_1\cup A_2$ 关于 λ 也是正集。

3. 证明当且仅当 $|\lambda|(M)=0$，$M\in\mathscr{F}$ 时，M 关于 λ 是零集。

4. 若 λ 是 $\mathscr{F}$ 上的广义测度，且 λ 只取有限值，证明

$$\lambda^+(E)=\sup\{\lambda(F):F\subset E,F\in\mathscr{F}\}$$
$$\lambda^-(E)=-\inf\{\lambda(F):F\subset E,F\in\mathscr{F}\}$$

5. 设 μ_1,μ_2 和 μ_3 都是 $(\Omega,\mathscr{F})$ 上的测度，证明 $\mu_1\ll\mu_1$，且由 $\mu_1\ll\mu_2$ 和 $\mu_2\ll\mu_3$ 可推出 $\mu_1\ll\mu_3$。举一个例子说明 $\mu_1\ll\mu_2$ 时不一定有 $\mu_2\ll\mu_1$。

6. 设 $\{\mu_n\}$ 是 $(\Omega,\mathscr{F})$ 的一个测度序列，$\mu_n(\Omega)\leqslant 1$，λ 定义为

$$\lambda(E)=\sum_{n=1}^{\infty}2^{-n}\mu_n(E),\quad \forall E\in\mathscr{F}$$

证明 λ 是测度且 $\mu_n\ll\lambda$，而且对一切 n 成立。

7. 设 λ 是 $(\Omega,\mathscr{F})$ 上的广义测度，μ 为其上的测度。若 $\lambda\ll\mu$，证明 $\lambda^+\ll\mu$，$\lambda^-\ll\mu$ 和 $|\lambda|\ll\mu$。

8. 设 λ,μ 是两个广义测度，其中 λ 是有限的，且 $\lambda\ll\mu$，证明对任给 $\varepsilon>0$，存在 $\delta>0$ 使得对满足 $|\mu(E)|(E)<\delta$ 的每一个 E，都有 $|\lambda|(E)<\varepsilon$。

9. 设 μ 由习题 1.8,8 定义，$E\subset N$，λ 定义为

$$\lambda(E)=\begin{cases}0, & \text{当 } E=\varnothing\\ +\infty, & \text{当 } E\neq\varnothing\end{cases}$$

若 $\mathscr{F}$ 为 N 的一切子集所成的 σ 代数，μ 为 $\mathscr{F}$ 上的有限测度，λ 为 $\mathscr{F}$ 上的无限测度，证明 $\lambda\ll\mu$，$\mu\ll\lambda$。

10. 利用习题 1.8,8 把 Radon-Nikodym 定理扩充到 μ 是 σ 有限测度，λ 是任意测度，$\lambda\ll\mu$ 的情形。这里 f 不一定取有限值。

11. (1) 设 Ω 是一不可数集，$\mathscr{F}$ 是由 E 为可数集或 $\Omega-E$ 为可数集的所有 E 集组成。定义

$$\mu(E)=\begin{cases}+\infty, & \text{若 } E \text{ 为可数集}\\ 0, & \text{否则}\end{cases}$$

而

$$\lambda(E)=\begin{cases}0, & E \text{ 为可数集}\\ +\infty, & E \text{ 为不可数集}\end{cases}$$

证明 $\lambda\ll\mu$，但 Radon-Nikodym 定理不成立。

(2) 设 $\Omega=[0,1]$，$\mathscr{F}$ 是 Ω 上的 Borel 子集所成类。μ 是 $\mathscr{F}$ 上的计数测度，λ 是 $\mathscr{F}$ 上 Lebesgue 测度，证明 λ 是有限测度且 $\lambda\ll\mu$，但 Radon-Nikodym 定理不成立。

12. 设 λ，μ 是 $(\Omega,\mathscr{F})$ 上 σ 有限测度，令 $\lambda\ll\mu$，$f=d\lambda/d\mu$，若 $g\in M^+(\Omega,\mathscr{F})$ 证明

$$\int g\,d\lambda=\int gf\,d\mu$$

[提示：先考虑简单函数再利用单调收敛定理]

13. 设 λ，μ，ν 是 $(\Omega,\mathscr{F})$ 上的 σ 有限测度。利用习题 1.8,12 证明：若 $\nu\ll\lambda$ 且 $\lambda\ll\mu$，那么

$$\frac{d\nu}{d\mu}=\frac{d\nu}{d\lambda}\frac{d\lambda}{d\mu}, \text{ a.e.}$$

若 $\lambda_i \ll \mu$，$i=1,2$，证明

$$\frac{d}{d\mu}(\lambda_1+\lambda_2)=\frac{d\lambda_1}{d\mu}+\frac{d\lambda_2}{d\mu}, \text{ a.e.}$$

14.若 λ 和 μ 是 σ 有限测度，$\lambda \ll \mu$ 且 $\mu \ll \lambda$，证明

$$\frac{d\lambda}{d\mu}=\frac{1}{d\mu/d\lambda}, \text{ a.e.}$$

15.若 λ 和 μ 为测度，$\lambda \ll \mu$ 且 $\lambda \perp \mu$，证明 $\lambda=0$。

16.若 λ 是广义测度（只取有限值），μ 是测度，证明 $\lambda \perp \mu$ 可推出 $\lambda^+ \perp \mu$，$\lambda^- \perp \mu$ 以及 $|\lambda| \perp \mu$。

17.证明 $(\Omega,\mathscr{F})$ 上所有取有限值的广义测度，在下列运算：

$$(c\mu)(E)=c\mu(E),\ (\lambda+\mu)(E)=\lambda(E)+\mu(E)$$

范数为

$$\|\mu\|=|\mu|(\Omega)$$

时组成 Banach 空间。

18.设定分布函数 $F(x)$ 在$[a,b]$上具有连续的导数，证明 $F(x)$ 在$[a,b]$上绝对连续。

1.9* 测度的生成

提要 本节讨论如何从代数(非 σ 代数)上的测度扩张到 σ 代数上的测度。本节定义了外测度，讨论了外测度的性质，并引进了 μ^*可测的概念。最后证明了全体 μ^*可测集组成一个 σ 代数，μ^* 是这个 σ 代数上的扩张测度。如果代数上的测度是 σ 有限的，那么扩张到 σ 代数上的测度是唯一的。

前面已经讨论过测度的定义、性质以及与它密切相关的积分等。本节着重讨论测度的构造。这一内容在一般测度论

著作中总是放在引入测度之前讲的，因为这是建立测度概念的基础。本书为了使读者先接触应用的内容，因此测度生成在这里只作一补充。

测度实际上是区间长度概念的推广，为了讨论的方便，我们以 Borel 集上的 Lebesgue 测度为重点来讲解。

半开区间$(a,b]$的长度是实数$b-a$，$(-\infty,b]$，$(a,+\infty)$和$(-\infty,+\infty)$的长度是$+\infty$。有限个两两不相交区间之和的长度定义为对应区间长度之和。如

$$E=\bigcup_{i=1}^{n}(a_i,b_i]$$

那么E的长度

$$l(E)=\sum_{i=1}^{n}(b_i-a_i)$$

下面我们先讨论在代数(非σ代数)上的测度。考虑下列形式的区间：

$(a,b]$，$(-\infty,b]$，$(a,+\infty)$，$(-\infty,+\infty)$及其有限联组成的集类 (1.9.1)

它们不是σ代数，而是代数。

定义1.9.1 设$\mathscr{A}$是Ω子集所成的类，若满足下列条件：

(1) $\varnothing$和 $\Omega\in\mathscr{A}$；

(2) 若 $E\in\mathscr{A}$，那么 $(\Omega-E)\in\mathscr{A}$；

(3) 若 $E_1,E_2,\cdots,E_n\in\mathscr{A}$，那么它们的有限联 $\bigcup_{i=1}^{n}E_i$ 也属于$\mathscr{A}$，则称$\mathscr{A}$是代数(并非σ代数)。

在代数上可以定义测度的概念。

定义 1.9.2 若$\mathscr{A}$是Ω子集所成的代数，μ是$\mathscr{A}$上广义实值函数，且满足下列条件：

(1) $\mu(\varnothing)=0$;

(2) $\mu(E)\geqslant 0$, $\forall\ E\in\mathscr{A}$;

(3) 若$\{E_n\}$是$\mathscr{A}$中两两不相交序列且$\bigcup\limits_{n=1}^{\infty}E_n\in\mathscr{F}$，那么

$$\mu\left(\bigcup_{n=1}^{\infty}E_n\right)=\sum_{n=1}^{\infty}\mu(E_n)$$

μ称为$\mathscr{A}$上的测度.

有理由认为长度就是测度，下面我们证明这一事实.

引理1.9.1 $\mathscr{A}$是式(1.9.1)形式有限联组成的集类，它是R中子集所成的代数，且长度是测度.

证明 容易看出$\mathscr{A}$是代数. 设l表示长度函数，那么定义 1.9.2 中 (1), (2) 不证自明. 为了证明长度满足定义1.9.2的(3)，只要证明式 (1.9.1) 形式的可列个集之联仍有式(1.9.1)的形式(这一点作为习题)且可列可加. 现在我们只证明可列可加性.

设对任何n，在区间$(a,b]$上作

$$a\leqslant a_1<b_1\leqslant a_2<\cdots<b_{n-1}\leqslant a_n<b_n\leqslant b$$

那么

$$\sum_{i=1}^{n}l((a_i,b_i])=\sum_{i=1}^{n}(b_i-a_i)$$
$$\leqslant b_n-a_1\leqslant b-a=l((a,b])$$

由于n的任意性，可推出

$$\sum_{i=1}^{\infty}l((a_i,b_i])\leqslant l((a,b]) \qquad (1.9.2)$$

反之，设$\varepsilon>0$，ε_i是一正数序列且满足$\sum\varepsilon_i<\varepsilon$. 令

$$I_1=(a_1-\varepsilon_1,b_1+\varepsilon_1)$$
$$I_i=(a_i,b_i+\varepsilon_i),\ i\geqslant 2$$

并设

$$(a,b]=\bigcup_{n=1}^{\infty}(a_i,b_i] \tag{1.9.3}$$

这样，开区间序列 $\{I_i;i\in N\}$ 覆盖了闭区间 $[a,b]$。由有限覆盖定理可知，存在有限个开区间 $I_1,I_2,\cdots,I_m$ 覆盖 $[a,b]$。令

$$a=a_1\leqslant a_2<b_1+\varepsilon_1<\cdots<a_m<b_{m-1}+\varepsilon_{m-1}$$
$$\leqslant b<b_m+\varepsilon_m$$

于是得

$$b-a\leqslant(b_m+\varepsilon_m)-a_1\leqslant\sum_{j=1}^{m}[(b_j+\varepsilon_j)-a_j]$$
$$<\sum_{j=1}^{m}(b_j-a_j)+\varepsilon\leqslant\sum_{j=1}^{\infty}(b_j-a_j)+\varepsilon$$

由 ε 的任意性可知

$$l((a,b])\leqslant\sum_{j=1}^{\infty}l((a_j,b_j])$$

把它与式(1.9.2)对照就可得出 l 在 $\mathscr{A}$ 上可列可加。 ▌

欲从代数上定义的测度推广到一般 σ 代数上的测度，需要引进外测度的概念。

定义 1.9.3 若 B 是 Ω 的任一子集，定义

$$\mu^*(B)=\inf_{\{E_j\}\subset\mathscr{A}}\sum_{j=1}^{\infty}\mu(E_j) \tag{1.9.4}$$

$$B\subset\bigcup_{j=1}^{\infty}E_j \tag{1.9.5}$$

那么 μ^* 称为由测度 μ 产生的外测度。外测度一般不是测度。

引理 1.9.2 由定义 1.9.3 定义的外测度 μ^* 具有下列性质：

(1) $\mu^*(\varnothing)=0$;

(2) $\mu^*(B)\geqslant 0$, $\forall B\subset\Omega$;

(3) 若 $A\subset B$ 则 $\mu^*(A)\leqslant\mu^*(B)$;

(4) 若 $B\in\mathscr{A}$ 则 $\mu^*(B)=\mu(B)$;

(5) 若 $\{B_n\}$ 是 Ω 子集序列，则

$$\mu^*\left(\bigcup_{n=1}^{\infty} B_n\right)\leqslant\sum_{n=1}^{\infty}\mu^*(B_n)$$

性质(3)称为外测度的单调性，性质(5)称为外测度的 σ 拟可加性.

证明 (1),(2),(3)是定义 1.9.3 的直接推论.

(4) 因为$\{B,\varnothing,\varnothing,\cdots\}$是 $\mathscr{A}$ 中的可列序列且其联等于 B，因此

$$\mu^*(B)\leqslant\mu(B)+0+\cdots=\mu(B)$$

反过来，若$\{E_n\}$是 $\mathscr{A}$ 中任一序列，$B\subset\bigcup E_n$ 那么 $B=\bigcup(B\cap E_n)$。因为 μ 是 $\mathscr{A}$ 上的测度，因此

$$\mu(B)\leqslant\sum_{n=1}^{\infty}\mu(B\cap E_n)\leqslant\sum_{n=1}^{\infty}\mu(E_n)$$

$$\mu(B)\leqslant\mu^*(B)$$

于是(4)得证.

为了证明(5)，设 $\varepsilon>0$ 是任给的正数，对于每一个 n 在 $\mathscr{A}$ 中选取一序列 $\{E_{n_k}\}$ 使得

$$B_n\subset\bigcup_{k=1}^{\infty}E_{n_k} \text{ 且 } \sum_{k=1}^{\infty}\mu(E_{n_k})\leqslant\mu^*(B_n)+\frac{\varepsilon}{2^n}$$

由于 $\{E_{n_k}\}$, n, $k\in N$ 是 $\mathscr{A}$ 中可列集，包含了联 $\bigcup B_n$. 由 μ^* 的定义可知

$$\mu^*\left(\bigcup_{n=1}^{\infty} B_n\right)\leqslant\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\mu(E_{n_k})\leqslant\sum_{n=1}^{\infty}\mu^*(B_n)+\varepsilon$$

由于 ε 的任意性，(5)得证. ∎

μ^* 虽然可以对 Ω 的所有子集定义，但 μ^* 没有可列可加性. 因此，我们要在包含 $\mathscr{A}$ 的 σ 代数上来考虑，使得 μ^* 具有可列可加性.

这个条件是由 Caratheodory 给出的.

定义 1.9.4 $E\subset\Omega$，若对于 Ω 的一切子集 A 有

$$\mu^*(A)=\mu^*(A\cap E)+\mu^*(A-E) \tag{1.9.6}$$

则称 E 是 μ^* 可测的.

全体 μ^* 可测集所成类记作 $\mathscr{A}^*$.

定理 1.9.1(Caratheodory扩张定理) **μ^* 可测集所成类 $\mathscr{A}^*$ 是包含 $\mathscr{A}$ 的 σ 代数. 若 $\{E_n\}$ 是 $\mathscr{A}^*$ 中的两两不相交集列，那么**

$$\mu^*\left(\bigcup_{n=1}^{\infty} E_n\right)=\sum_{n=1}^{\infty}\mu^*(E_n) \tag{1.9.7}$$

证明 显然 $\varnothing$ 和 Ω 是 μ^* 可测集. 若 $E\in\mathscr{A}^*$，那么其补 $\Omega-E\in\mathscr{A}^*$.

下面我们证明 $\mathscr{A}^*$ 关于交运算封闭. 假设 E 和 F 都是 μ^* 可测集. 那么对任何 $A\subset\Omega$，有

$$\mu^*(A\cap F)=\mu^*(A\cap F\cap E)+\mu^*((A\cap F)-E)$$

因为 $F\in\mathscr{A}^*$，所以

$$\mu^*(A)=\mu^*(A\cap F)+\mu^*(A-F)$$

令 $B=A-(E\cap F)$，容易看到 $B\cap F=(A\cap F)-E$ 和 $B-F=A-F$；又因为 $F\in\mathscr{A}^*$，所以

$$\mu^*(A-(E\cap F))=\mu^*((A\cap F)-E)+\mu^*(A-F)$$

综合上面三个式子得

$$\mu^*(A)=\mu^*(A\cap E\cap F)+\mu^*(A-(E\cap F))$$

于是 $E\cap F\in\mathscr{A}^*$. 由于 $\mathscr{A}^*$ 关于交与补是封闭的，因此 $\mathscr{A}^*$ 是代数.

现在，我们要证明 $\mathscr{A}^*$ 是 σ 代数，即 μ^* 在 $\mathscr{A}^*$ 上可列

可加. 设 $\{E_k\}$ 是 $\mathscr{A}^*$ 中两两不相交序列，$E = \bigcup E_k$. 由上面的证明可以知道

$$F_n = \bigcup_{k=1}^{n} E_k \in \mathscr{A}^*$$

若 $A \subset \Omega$，那么

$$\begin{aligned}\mu^*(A) &= \mu^*(A \cap F_n) + \mu^*(A - F_n) \\ &= \sum_{k=1}^{n} \mu^*(A \cap E_k) + \mu^*(A - F_n)\end{aligned}$$

因为 $F_n \subset E$，所以 $A - E \subset A - F_n$. 令 $n \to \infty$，由上面的不等式可得

$$\sum_{k=1}^{\infty} \mu^*(A \cap E_k) + \mu^*(A - E) \leqslant \mu^*(A)$$

另一方面，由引理 1.9.2(5) 有

$$\mu^*(A \cap E) \leqslant \sum_{k=1}^{\infty} \mu^*(A \cap E_k)$$

$$\mu^*(A) \leqslant \mu^*(A \cap E) + \mu^*(A - E)$$

把最后三个不等式联系起来可得

$$\begin{aligned}\mu^*(A) &= \mu^*(A \cap E) + \mu^*(A - E) \\ &= \sum_{k=1}^{\infty} \mu^*(A \cap E_k) + \mu^*(A - E)\end{aligned}$$

这就证明了 $E = \bigcup\limits_{k=1}^{\infty} E_k$ 是 μ^* 可测的. 取 $A = E$ 可以得到式(1.9.7).

剩下的是要证明 $\mathscr{A} \subset \mathscr{A}^*$. 在引理 1.9.2(4) 中已经证明，若 $E \in \mathscr{A}$ 则 $\mu(E) = \mu^*(E)$，现在我们需要证明 $E \in \mathscr{A}^*$，即 E 是 μ^* 可测的. 设 A 是 Ω 的任一子集，由引理 1.9.2(5)得

$$\mu^*(A) \leqslant \mu^*(A \cap E) + \mu^*(A - E)$$

为了建立反向不等式，设任给 $\varepsilon>0$，$\{F_n\}$ 是 $\mathscr{A}$ 中的序列，$A\subset\bigcup F_n$ 且

$$\sum_{n=1}^{\infty}\mu(F_n)\leqslant\mu^*(A)+\varepsilon$$

因为 $A\cup E\subset(F_n\cap E)$ 且 $A-E\subset\bigcup(F_n-E)$，由引理 1.9.2 (5) 得

$$\mu^*(A\cap E)\leqslant\sum_{n=1}^{\infty}\mu(F_n\cap E)$$

$$\mu^*(A-E)\leqslant\sum_{n=1}^{\infty}\mu(F_n-E)$$

因此有

$$\mu^*(A\cap E)+\mu^*(A-E)\leqslant\sum_{n=1}^{\infty}\{\mu(F_n\cap E)+\mu(F_n-E)\}$$

$$=\sum_{n=1}^{\infty}\mu(F_n)\leqslant\mu^*(A)+\varepsilon$$

由 ε 的任意性，可知

$$\mu^*(A\cap E)+\mu^*(A-E)\leqslant\mu^*(A)$$

由此得

$$\mu^*(A)=\mu^*(A\cap E)+\mu^*(A-E)$$

即 E 是 μ^* 可测集，故 $\mathscr{A}\subset\mathscr{A}^*$.

Caratheodory 扩张定理说明：$\mathscr{A}$ 上的测度 μ 总可以扩张到 $\mathscr{A}^*$ 上的测度 μ^*，且 $\mathscr{A}^*\supset\mathscr{A}$. ∎

若 $E\in\mathscr{A}^*$，$\mu(E)=0$，$B\subset E$，那么 $B\in\mathscr{A}^*$ 且 $\mu(B)=0$，这时 $\mathscr{A}^*$ 称为完备 σ 代数. 容易证明 $\mathscr{A}^*$ 是完备的.

为了证明这一点，设 A 是 Ω 的任一子集，由引理 1.9.2 (3) 得

$$\mu^*(A) = \mu^*(E) + \mu^*(A)$$
$$\geqslant \mu^*(A \cap B) + \mu^*(A - B)$$

由引理 1.9.2(5) 得

$$\mu^*(A) \leqslant \mu^*(A \cap B) + \mu^*(A - B)$$

因此 B 是 μ^* 可测集且

$$0 \leqslant \mu^*(B) \leqslant \mu^*(E) = 0$$ ▮

下面可以证明若 μ 是 σ 有限测度，则可以唯一地扩张为 $\mathscr{A}^*$ 上的测度.

定理 1.9.2 (Hahn 扩张定理) 设 μ 是代数 $\mathscr{A}$ 上的 σ 有限测度，那么存在 $\mathscr{A}^*$ 的唯一扩张测度.

证明 定理 1.9.1 证明了即使没有 σ 有限条件 μ^* 仍是 $\mathscr{A}^*$ 上测度. 为了证明唯一性，设 ν 也是 $\mathscr{A}^*$ 的测度且在 $\mathscr{A}$ 上与 μ 一致，即 ν 也是 μ 在 $\mathscr{A}^*$ 上的扩张测度.

先假定 μ, μ^*, ν 都是有限测度.

$$E \in \mathscr{A}^*, \quad E_n \in \mathscr{A}, \quad n = 1, 2, \cdots \text{且} \quad E \subset \bigcup E_n$$

因为 ν 在 $\mathscr{A}$ 上与 μ 重合，所以

$$\nu(E) \leqslant \nu\left(\bigcup_{n=1}^{\infty} E_n\right) \leqslant \sum_{n=1}^{\infty} \nu(E_n) = \sum_{n=1}^{\infty} \mu(E_n)$$

故

$$\nu(E) \leqslant \mu^*(E), \quad \forall E \in \mathscr{A}^*$$

因为 μ^* 与 ν 都是可加的，所以

$$\mu^*(E) + \mu^*(\Omega - E) = \nu(E) + \nu(\Omega - E)$$
$$\mu^*(\Omega - E) \geqslant \nu(\Omega - E)$$

故

$$\mu^*(E) \leqslant \nu(E), \quad \forall E \in \mathscr{A}^*$$
$$\mu^*(E) = \nu(E), \quad \forall E \in \mathscr{A}^*$$

因此 $\mu^* = \nu$，即当 μ 是有限测度时扩张是唯一的.

假设 μ 是 σ 有限测度，$\{F_n\}$ 是 $\mathscr{A}$ 中的增序列且

$\mu(F_n)<+\infty$，$\Omega=\bigcup F_n$．由上面的证明有

$$\mu^*(E\cap F_n)=\nu(E\cap F_n),\quad \forall E\in\mathscr{A}^*$$

所以

$$\begin{aligned}\mu^*(E)&=\lim\mu^*(E\cap F_n)\\&=\lim\nu(E\cap F_n)=\nu(E)\end{aligned}$$

即 μ^* 与 ν 在 $\mathscr{A}^*$ 重合． ∎

从上面的讨论可知，测度的构造如下：先在代数上定义测度，然后扩张到包含 $\mathscr{A}$ 的 $\mathscr{A}^*$ 上的测度 μ^*．$\mathscr{A}^*$ 是一个完备 σ 代数．

习　题　1.9

1．证明形如式(1.9.1)集的全体成一个代数．

2．证明形如 (a,b)，$(-\infty,b)$，$(a,+\infty)$，$(-\infty,+\infty)$ 有限联的全体 G 不是 R 上的代数．然而由 G 产生的 σ 代数是 Borel 代数．

3．若 $(a,+\infty)$ 是不相交列 $(a_n,b_n]$ 的联，证明

$$\sum_{n=1}^{\infty}l((a_n,b_n])=+\infty$$

4．设 Ω 是满足 $0<r\leqslant1$ 的全体有理数，$\mathscr{A}$ 是形如 $\{r\in\Omega:a<r\leqslant b\}$，$0\leqslant a\leqslant b\leqslant1$，$a,b\in\Omega$ 的"半开区间"的有限联全体．证明 $\mathscr{A}$ 是代数．

5．若 E 是 R 的可数子集，证明 E 的 Lebesgue 测度为零．

6．设 $I_n=(n,n+1]$，$n=0,\pm1,\pm2,\cdots$．若 $E\subset\bigcup\limits_{i=1}^{n}I_i$（有限联），那么 $l^*(E)<+\infty$．作一 Lebesgue 可测集 E，$l^*(E)<+\infty$，使得 $l^*(E\cap I_n)>0$，对一切 n 成立．证明当且仅当 $E\cap I_n$，$\forall n=1,2,\cdots$ 是 Lebesgue 可测集时 E 是 Lebesgue 可测集．

7．若 A 是 R 的 Lebesgue 可测子集且 $\varepsilon>0$，证明存在一个开集 G_ε，$G_\varepsilon\supset A$，使得

$$l^*(A)\leqslant l^*(G_\varepsilon)\leqslant l^*(A)+\varepsilon$$

8．设 B 是 R 的一个 Lebesgue 可测子集且 $\varepsilon>0$．若 $B\subset I_n=(n,$

$n+1]$，证明存在一个紧集 $K_\varepsilon \subset B$ 使得

$$l^*(K_\varepsilon) \leqslant l^*(B) \leqslant l^*(K_\varepsilon) + \varepsilon$$

[提示：应用习题1.9，7，$A = I_n - B$]

9．若 A 是 R 中任一 Lebesgue 可测集，利用上面的习题证明

$$l^*(A) = \inf\{l^*(G): A \subset G, G \text{ 为开集}\}$$

$$l^*(A) = \sup\{l^*(K): K \subset A, A \text{ 为紧集}\}$$

10．令 $\lambda = l^*$ 表示 R 上的 Lebesgue 测度，设 A 是一个 Lebesgue 可测集且 $\lambda(A) < +\infty$．证明对 $\varepsilon > 0$，存在一开集它是有限个开区间之联且使得

$$\|\chi_A - \chi_G\|_1 = |\lambda(A) - \lambda(G)| < \varepsilon$$

进一步证明，对 $\varepsilon > 0$ 存在一连续函数 f，使得

$$\|\chi_A - f\|_1 = \int |\chi_A - f| d\lambda < \varepsilon$$

1.10 乘积测度

提要 本节在引入乘积空间，可测矩形和乘积 σ 代数后，讨论了乘积测度与重积分．乘积测度定理保证了乘积 σ 代数上测度的存在性． 单调类引理是证明 Tonelli 定理的工具之一．本节最后证明了累次积分可交换的条件是函数关于乘积测度可积——这就是著名的 Fubini 定理．

为了研究多维随机现象，我们必须研究乘积空间和乘积测度，以及由此派生出来的重积分．这些内容不仅对于概率论的研究是重要的，而且对于测度论本身也是基本的．

设 Ω_1 和 Ω_2 是两个集，Ω_1 与 Ω_2 的笛卡尔乘积 $\Omega_1 \times \Omega_2$ 是一切有序点偶 (ω_1, ω_2) $(\omega_1 \in \Omega_1, \omega_2 \in \Omega_2)$ 的全体所组成的集，即

$$\Omega_1 \times \Omega_2 = \{(\omega_1,\omega_2):\ \omega_1 \in \Omega_1,\ \omega_2 \in \Omega_2\}$$

如果 Ω_1 与 Ω_2 中有一个是空集，则定义 $\Omega_1 \times \Omega_2 = \varnothing$。

积集 $\Omega_1 \times \Omega_2$ 也称为乘积空间。

定义 1.10.1 若 $(\Omega,\ \mathscr{F}_1)$ 和 $(\Omega_2,\ \mathscr{F}_2)$ 都是可测空间，$A \in \mathscr{F}_1, B \in \mathscr{F}_2$，则 $A \times B$ 称为 $Z = \Omega_1 \times \Omega_2$ 中的可测矩形．用 Z_0 表示所有可测矩形组成的集类。

在习题 1.10,4 中，我们将证明 Z_0 中任一集可以表示为 Z 中两两不相交的有限个可测矩形之联。

引理 1.10.1 集类 Z_0 是 Z 子集所成的代数。

证明 显然 Z_0 中有限集之联仍属于 Z_0。根据习题 1.10,5 的第一部分，可测矩形之补属于 Z_0。利用 De Morgan 定律，Z_0 中任何集之补属于 Z_0. ∎

定义 1.10.2 若 $(\Omega_1, \mathscr{F}_1)$ 和 $(\Omega_2, \mathscr{F}_2)$ 为可测空间，由可测矩形 $A \times B$，$A \in \mathscr{F}_1$，$B \in \mathscr{F}_2$ 生成的 σ 代数，记作 $\mathscr{F} = \mathscr{F}_1 \times \mathscr{F}_2$。$\mathscr{F}$ 中的集称为可测集或 Z 的可测子集。

$(\Omega_1, \mathscr{F}_1, \mu)$ 和 $(\Omega_2, \mathscr{F}_2, \nu)$ 是测度空间，我们在 $\mathscr{F} = \mathscr{F}_1 \times \mathscr{F}_2$ 上定义一个测度 π，它是 μ 与 ν 的乘积，即

$$\pi(A \times B) = \mu(A)\nu(B),\ \forall A \in \mathscr{F}_1,\ B \in \mathscr{F}_2$$

这里我们仍约定 $0 \cdot (\pm\infty) = 0$。

定理 1.10.1（乘积测度定理） 若 $(\Omega_1, \mathscr{F}_1, \mu)$ 和 $(\Omega_2, \mathscr{F}_2, \nu)$ 是两个测度空间，那么存在定义于 $\mathscr{F} = \mathscr{F}_1 \times \mathscr{F}_2$ 上的测度 π，使得

$$\pi(A \times B) = \mu(A)\nu(B),\ \forall A \in \mathscr{F}_1, B \in \mathscr{F}_2 \quad (1.10.1)$$

若这些测度空间是 σ 有限的，那么测度 π 是唯一的。

证明 若 $A \times B$ 是两两不相交可测矩阵列 $\{A_j \times B_j\}$ 之联

$$\chi_{A \times B}(\omega_1, \omega_2) = \chi_A(\omega_1)\chi_B(\omega_2)$$

$$= \sum_{j=1}^{\infty} \chi_{A_j}(\omega_1)\chi_{B_j}(\omega_2), \forall A \in \mathscr{F}_1, B \in \mathscr{F}_2$$

固定 ω_1，对 ν 积分并利用单调收敛定理，可得

$$\chi_A(\omega_1)\nu(B) = \sum_{j=1}^{\infty} \chi_{A_j}(\omega_1)\nu(B_j)$$

对 μ 积分再利用单调收敛定理，得

$$\mu(A)\nu(B) = \sum_{j=1}^{\infty} \mu(A_j)\nu(B_j)$$

设 $E \in Z_0$，不失一般性，我们可以假定

$$E = \bigcup_{j=1}^{n} (A_j \times B_j)$$

其中 $A_j \times B_j$ 是两两不相交的可测矩形。 如果我们定义 $\pi_0(E)$ 为

$$\pi_0(E) = \sum_{j=1}^{n} \mu(A_j)\nu(B_j)$$

根据 1.9 节的讨论，π_0 在 Z_0 上可列可加. 由定理 1.9.1 知存在 π_0 的扩张测度 π，它是由 Z_0 生成的 σ 代数 $\mathscr{F}$ 上的测度。因为 π 是 π_0 的扩张测度，因此式(1.10.1)成立.

如果 $(\Omega_1, \mathscr{F}, \mu)$ 和 $(\Omega_2, \mathscr{F}_2, \nu)$ 都是 σ 有限测度空间，那么 π_0 是 Z_0 上的 σ 有限测度. 由 Hahn 扩张定理，满足式(1.10.1)的扩张测度 π 是唯一的. ▮

定理 1.10.1 说明了在由可测矩形产生的 σ 代数 $\mathscr{F}$ 上，满足式 (1.10.1) 的测度的存在性. 这种测度 π 称为乘积测度. 若 μ 和 ν 均 σ 有限，那么它们必有唯一的乘积测度。1.9 节讨论的扩张定理保证了这种测度的唯一性。然而，如果 μ 和 ν 并非 σ 有限的话，在习题中可以看到，Z 上可以有两个不同的测度。

为了讨论关于乘积测度的积分，我们先引进下列概念。

定义 1.10.3 若 $E\subset Z=\Omega_1\times\Omega_2$ 且 $\omega_1\in\Omega_1$ 那么E的ω_1截口定义为

$$E_{\omega_1}=\{\omega_2\in\Omega_2:(\omega_1,\omega_2)\in E\}$$

同样E的ω_2截口为

$$E_{\omega_2}=\{\omega_1\in\Omega_1:(\omega_1,\omega_2)\in E\}$$

若f是定义在Z上取值于$\bar{R}=[-\infty,+\infty]$的函数，于是$f$的$\omega_1$截口是$\Omega_2$上的函数，即

$$f_{\omega_1}(\omega_2)=f(\omega_1,\omega_2),\quad \omega_2\in\Omega_2$$

f的ω_2截口f_{ω_2}是Ω_1上的函数

$$f_{\omega_2}=f(\omega_1,\omega_2),\quad \omega_1\in\Omega_1$$

引理 1.10.2 若E是Z的可测子集，那么

(1) E的每一个截口可测.

(2) 若f是定义在Z上取值于$\bar{R}$的可测函数，那么f的每一截口可测.

证明 (1) 若 $E=A\times B$ 且 $\omega_1\in\Omega_1$，那么E的ω_1截口为

$$E_{\omega_1}=\begin{cases}B, & \forall\omega_1\in A\\ \varnothing, & \omega_1\notin A\end{cases}$$

设$\mathscr{E}$是$\mathscr{F}$中每个截口均可测的集所成的类，那么可测矩形$A\times B\in\mathscr{E}$.

容易证明实际上$\mathscr{E}=\mathscr{F}$（习题 1.10,9）

(2) 设 $\omega_1\in\Omega$，$\alpha\in R$，那么

$$\begin{aligned}&\{\omega_2\in\Omega_2:\ f_{\omega_1}(\omega_2)>\alpha\}\\ &=\{\omega_2\in\Omega_2:\ f(\omega_1\omega_2)>\alpha\}\\ &=\{(\omega_1,\omega_2)\in\Omega_1\times\Omega_2:\ f(\omega_1\omega_2)>\alpha\}_{\omega_1}\end{aligned}$$

因为f是$\mathscr{F}$可测的，所以f_{ω_1}关于$\mathscr{F}_2$可测. 同理可证明f_{ω_2}关于$\mathscr{F}_1$可测. ∎

设$\mathscr{M}$是空间Ω（或Z)的某些子集所成的类，如果对于

$\mathscr{M}$ 中的任何一个单调集列 $\{E_n\}$，都有

$$\lim E_n \in \mathscr{M}$$

则称 $\mathscr{M}$ 为单调类。

设 $\mathscr{A}$ 是 Ω（或 Z）的某些子集所成的非空类，S 是包含 $\mathscr{A}$ 的最小 σ 代数，称为由 $\mathscr{A}$ 生成的 σ 代数，记作 $S(\mathscr{A})$。$\mathscr{M}$ 是包含 $\mathscr{A}$ 的最小单调类，称为由 $\mathscr{A}$ 生成的单调类，记作 $\mathscr{M}(\mathscr{A})$。

引理 1.10.3 σ 代数必为单调类。

证明 设 $\mathscr{F}$ 是 σ 代数，$\{A_n\}$ 是一单调集列，且 $A_n \in \mathscr{F}$，由于 $\mathscr{F}$ 关于可列和运算封闭，故当 $\{A_n\}$ 单调增加时有

$$\lim A_n = \bigcup_{n=1}^{\infty} A_n \in \mathscr{F}$$

由于 $\mathscr{F}$ 关于可列交运算封闭，故当 $\{A_n\}$ 单调减小时，有

$$\lim A_n = \bigcap_{n=1}^{\infty} A_n \in \mathscr{F}$$

这就证明了 $\mathscr{F}$ 是单调类。 ∎

$S(\mathscr{A})$ 是包含 $\mathscr{A}$ 的 σ 代数，由引理 1.10.3 知它是单调类，而 $\mathscr{M}(\mathscr{A})$ 是包含 $\mathscr{A}$ 的最小单调类，故

$$\mathscr{M}(\mathscr{A}) \subset S(\mathscr{A})$$

引理 1.10.4 若 $\mathscr{A}$ 本身是代数，那么

$$\mathscr{M}(\mathscr{A}) = S(\mathscr{A})$$

证明 上面已经说过 $\mathscr{M}(\mathscr{A}) \subset S(\mathscr{A})$。为了证明反向的包含关系，只要证明 $\mathscr{M}(\mathscr{A})$ 是一个代数就足够了。因为若 $\mathscr{M}(\mathscr{A})$ 是代数又是单调类，那么它必定是 σ 代数。（这一点容易验证，我们将它留作习题）

对于任意一个集 $E \in \mathscr{M}$，记

$\kappa(E) = \{F \mid F \in \mathscr{M}, F - E, E - F, E \cup F \in \mathscr{M}\}$

设 $\{F_n\}$ 是 $\kappa(E)$ 中任一单调集列，则

$$E-F_n\in\mathscr{M},\ F_n-E\in\mathscr{M},\ F_n\cup E\in\mathscr{M}$$

显然集列 $\{E-F_n\},\{F_n-E\},\{F_n\cup E\}$ 都是单调的，因为 $\mathscr{M}$ 是单调类，故上述三个序列的极限都属于 $\mathscr{M}$，但

$$\lim_{n\to\infty}(E-F_n)=E-\lim_{n\to\infty}F_n,\ \in\mathscr{M}$$

$$\lim_{n\to\infty}(F_n-E)=\lim_{n\to\infty}F_n-E,\ \in\mathscr{M}$$

$$\lim_{n\to\infty}(E\cup F_n)=E\cup\lim_{n\to\infty}F_n,\ \in\mathscr{M}$$

所以 $\lim_{n\to\infty}F_n\in\kappa(E)$，这说明 $\kappa(E)$ 是单调类。

现在证明 $\mathscr{M}(\mathscr{A})$ 是代数。记

$$\kappa_1(A)=\{B\mid B\subset\Omega\ \text{且}\ A-B,B-A,A\cup B\ \text{都属于}\ \mathscr{M}(\mathscr{A})\}$$

由于 $\kappa_1(A)$ 的定义中 A 和 B 的地位是对称的，故 $B\in\kappa_1(A)$ 与 $A\in\kappa_1(B)$ 等价。由于 $\mathscr{A}$ 是代数，故对于任何 $A\in\mathscr{A}$，$B\in\mathscr{A}$ 必有 $B\in\kappa_1(A)$，故 $\mathscr{A}\subset\kappa_1(A)$。上面已经证明 $\kappa_1(A)$ 是单调类，故

$$\mathscr{M}(\mathscr{A})\subset\kappa_1(A)$$

即对于 $A\in\mathscr{A}$，$B\in\mathscr{M}(\mathscr{A})$，有

$$B\in\kappa_1(\mathscr{A}),\ \text{即}\ A\in\kappa_1(B)$$

因此当 $B\in\mathscr{M}(\mathscr{A})$ 时，有

$$\mathscr{A}\subset\kappa_1(B)$$

由于 $\kappa_1(B)$ 是单调类，故得 $\mathscr{M}(\mathscr{A})\subset\kappa_1(B)$，这就是说，对于任何 $A\in\mathscr{M}(\mathscr{A})$，$B\in\mathscr{M}(\mathscr{A})$ 都有 $A\in\kappa_1(B)$。由 $\kappa_1(B)$ 的定义知 $A-B,B-A,A\cup B\in\mathscr{M}(\mathscr{A})$，且 $\Omega\in\mathscr{M}(\mathscr{A})$。故 $\mathscr{M}(\mathscr{A})$ 是代数。 ▌

由单调类引理知，包含了代数 $\mathscr{A}$ 的单调类必包含由 $\mathscr{A}$ 生成的 σ 代数。

引理 1.10.5 设 $(\Omega_1,\mathscr{F}_1,\mu)$ 与 $(\Omega_2,\mathscr{F}_2,\nu)$ 都是 σ

有限测度空间，$E\in\mathscr{F}=\mathscr{F}_1\times\mathscr{F}_2$，那么

$$f(\omega_1)=\nu(E_{\omega_1}),\quad g(\omega_2)=\mu(E_{\omega_2}) \tag{1.10.2}$$

是可测函数，且

$$\int_{\Omega_1} f d\mu=\pi(E)=\int_{\Omega_2} g d\nu \tag{1.10.3}$$

证明 首先我们假定测度空间是有限的，并令

$$\boldsymbol{M}=\left\{E\in\mathscr{F}:\int_{\Omega_1} f d\mu=\pi(E)=\int_{\Omega_2} g d\nu\right\}$$

我们将证明 $\boldsymbol{M}=\mathscr{F}$，这只要证明 $\boldsymbol{M}$ 是包含代数 Z_0的单调类．事实上，若 $E=A\times B$，$A\in\mathscr{F}_1$，$B\in\mathscr{F}_2$，那么

$$f(\omega_1)=\chi_A(\omega_1)\nu(B),\quad g(\omega_2)=\chi_B(\omega_2)\mu(A)$$

$$\int_{\Omega_1} f d\mu=\mu(A)\nu(B)=\int_{\Omega_2} g d\nu$$

因为 Z_0 中任一个元素可以写成有限个不相交矩阵的联，所以 $Z_0\subset\boldsymbol{M}$

现在证明 $\boldsymbol{M}$ 是单调类．设 $\{E_n\}$ 是 $\boldsymbol{M}$ 中一个增序列，$\bigcup E_n=E$．所以

$$f_n(\omega_1)=\nu((E_n)_{\omega_1}),\quad g_n(\omega_2)=\mu((E_n)_{\omega_2})$$

是可测函数且

$$\int_{\Omega_1} f_n d\mu=\pi(E_n)=\int_{\Omega_2} g_n d\nu$$

$\{f_n\}$，$\{g_n\}$ 是两个单调增序列，且分别收敛于

$$f(\omega_1)=\nu(E_{\omega_1}),\quad g(\omega_2)=\mu(E_{\omega_2})$$

π 是测度，利用单调收敛定理，得

$$\int_{\Omega_1} f d\mu=\pi(E)=\int_{\Omega_2} g d\nu$$

所以 $E\in\boldsymbol{M}$．同理可以证明，若 $\{F_n\}$ 是 $\boldsymbol{M}$ 中的递减序列，$F=\bigcap F_n$，$F\in\boldsymbol{M}$，则 $\boldsymbol{M}$ 是单调类，由单调类引理得 $\boldsymbol{M}=\mathscr{F}$．

如果测度空间是 σ 有限的，$\Omega=\Omega_1\times\Omega_2$，$\Omega$ 是增序列 $\{Z_n\}$ 之和且 $\mu(Z_n)<+\infty$。利用上面的讨论和单调收敛定理得

$$\int_{\Omega_1\cap Z_n} f d\mu=\pi(E\cap Z_n)=\int_{\Omega_2\cap Z_n} g d\nu$$

由积分及测度的可加性引理得证。 ∎

定理 1.10.2（Tonelli 定理） 设 $(\Omega_1,\mathscr{F}_1,\mu)$ 和 $(\Omega_2,\mathscr{F}_2,\nu)$ 是 σ 有限测度空间，ϕ 是 $\Omega=\Omega_1\times\Omega_2$ 上取值于 $\bar{R}$ 的非负可测函数。那么下列函数：

$$f(\omega_1)=\int_{\Omega_2}\phi_{\omega_1}d\nu,\ g(\omega_2)=\int_{\Omega_1}\phi_{\omega_2}d\mu \qquad (1.10.4)$$

是可测函数且

$$\int_{\Omega_1} f d\mu=\int_{\Omega}\phi d\pi=\int_{\Omega_2} g d\nu \qquad (1.10.5)$$

换句话说

$$\int_{\Omega_1}\left(\int_{\Omega_2}\phi d\nu\right)d\mu=\int_{\Omega}\phi d\pi=\int_{\Omega_2}\left(\int_{\Omega_1}\phi d\mu\right)d\nu \qquad (1.10.6)$$

证明 若 ϕ 是 $\mathscr{F}$ 中集 E 的示性函数，由引理 1.10.5 可知，命题成立。由于积分具有线性，因此本命题对于简单可测函数成立。

若 ϕ 是定义于 Ω，取值于 $\bar{R}$ 的任意非负可测函数。由引理 1.2.5 知，存在一非负简单可测函数列 ϕ_n 单调增地收敛于 ϕ。若

$$\varphi_n(\omega_1)=\int_{\Omega_2}(\phi_n)_{\omega_1}d\nu,\quad \psi_n(\omega_2)=\int_{\Omega_1}(\phi_n)_{\omega_2}d\mu \qquad (1.10.7)$$

则 φ_n，ψ_n 是单调增且可测的函数列，$\varphi_n\to f$，$\psi_n\to g$，由单调收敛定理得

$$\int_{\Omega_1} f d\mu=\lim\int_{\Omega_1}\varphi_n d\mu=\lim\int_{\Omega}\phi_n d\pi$$

$$= \lim \int_{\Omega_2} \phi_n d\nu = \int_{\Omega_2} g d\nu$$

同理

$$\int_{\Omega} \phi d\pi = \lim \int_{\Omega} \phi_n d\pi$$

式(1.10.5)得证。 ∎

在习题中，我们将看到，如果 ϕ 不是非负函数或 μ,ν 不是 σ 有限，那么 Tonelli 定理不一定成立。

定理 1.10.3 (Fubini 定理) 设 $(\Omega_1, \mathscr{F}_1, \mu)$ 和 $(\Omega_2, \mathscr{F}_2, \nu)$ 是 σ 有限测度空间，π 是 $\mathscr{F} = \mathscr{F}_1 \times \mathscr{F}_2$ 上 μ 与 ν 的乘积测度，若 $\Omega = \Omega_1 \times \Omega_2$ 上定义的实函数 ϕ，关于 π 可积，那么存在广义实函数

$$f(\omega_1) = \int_{\Omega_2} \phi_{\omega_1} d\nu, \qquad g(\omega_2) = \int_{\Omega_1} \phi_{\omega_2} d\mu \qquad (1.10.8)$$

有有限积分且

$$\int_{\Omega_1} f d\mu = \int_{\Omega} \phi d\pi = \int_{\Omega_2} g d\nu \qquad (1.10.9)$$

换言之

$$\int_{\Omega_1} \left(\int_{\Omega_2} \phi d\nu \right) d\mu = \int_{\Omega} \phi d\pi = \int_{\Omega_2} \left(\int_{\Omega_1} \phi d\mu \right) d\nu \quad (1.10.10)$$

证明 因为 ϕ 关于 π 可积，ϕ 的正部 ϕ^+ 与负部 ϕ^- 也可积。对于 ϕ^+ 利用 Tonelli 定理得

$$\int_{\Omega_1} f^+ d\mu = \int_{\Omega} \phi^+ d\pi = \int_{\Omega_2} g^+ d\nu$$

对于 ϕ^- 有

$$\int_{\Omega_1} f^- d\mu = \int_{\Omega} \phi^- d\pi = \int_{\Omega_2} g^- d\nu$$

因为这些积分均有限，故

$$\int_{\Omega_1} f d\mu = \int_{\Omega} \phi d\pi = \int_{\Omega_2} g d\nu$$ ∎

从习题中可以看到，如果ϕ不可积，Fubini 定理可能不成立。

习 题 1.10

1.设 $A\subset\Omega_1$，$B\subset\Omega_2$. 若A或B是空集，证明 $A\times B=\varnothing$. 反过来，若 $A\times B=\varnothing$，那么$A=\varnothing$或 $B=\varnothing$。

2.若 $A_j\subset\Omega_1$，$B_j\subset\Omega_2$，$\forall j=1,2$，且 $A_1\times B_1=A_2\times B_2$，证明 $A_1=A_2$ 且 $B_1=B_2$。

3.若 $A_j\subset\Omega_1$，$B_j\subset\Omega_2$，$\forall j=1,2$，证明

$$(A_1\times B_1)\cup(A_2\times B_2)=[(A_1-A_2)\times B_1]$$
$$\cup[(A_1\cap A_2)\times(B_1\cup B_2)]\cup[(A_2-A_1)\times B_2]$$

等式右边的集均不相交.

4.设 $(\Omega_1,\mathscr{F}_1)$ 和 $(\Omega_2,\mathscr{F}_2)$ 都是可测空间. 若 $A_j\in\mathscr{F}_1$，$B_j\in\mathscr{F}_2$，$\forall j=1,2,\cdots,n$，证明

$$\bigcup_{j=1}^{n}(A_j\times B_j)$$

可以写成 $\Omega=\Omega_1\times\Omega_2$ 中有限个两两不相交的矩形之和。

5.设 $A_j\subset\Omega_1$，$B_j\subset\Omega_2$，$\forall j=1,2$，证明

$$(A_1\times B_1)-(A_2\times B_2)=[(A_1\cap A_2)\times(B_1-B_2)]$$
$$\cup[(A_1-A_2)\times B_1]$$
$$(A_1\times B_1)\cap(A_2\times B_2)=(A_1\cap A_2)\times(B_1\cap B_2)$$

6.设 $R=(-\infty,\infty)$，$\mathscr{B}$是 Borel 代数，$(R,\mathscr{B})$ 是可测空间，证明$R\times R$中每一个开集都属于 $\mathscr{B}\times\mathscr{B}$，事实上，$\mathscr{B}\times\mathscr{B}$ 是由$R\times R$中开集生成的σ代数(换句话说，$\mathscr{B}\times\mathscr{B}$是$R\times R$上的 Borel 代数)。

7. f 和 g 分别是 Ω_1 和 Ω_2 上的实函数. 假如f是$\mathscr{F}_1$可测的，g是$\mathscr{F}_2$可测的，且

$$h(\omega_1,\omega_2)=f(\omega_1)g(\omega_2)$$

证明h是 $\mathscr{F}=\mathscr{F}_1\times\mathscr{F}_2$ 可测的.

8. E是R的一个子集，令

$$\gamma(E)=\{(x,y)\in R\times R;\ x-y\in E\}$$

若 $E\in\mathscr{B}$，证明 $\gamma(E)\in\mathscr{B}\times\mathscr{B}$. 利用这一点再证明，若 $f(x)$ 是R上

的 Borel 可测函数，那么 $F(x,y)=f(x-y)$ 关于 $\mathscr{B}\times\mathscr{B}$ 可测。

9. 设 E 和 F 是 $Z=X\times Y$ 的子集，$x\in X$，证明 $(E-F)_x=E_x-F_x$；若 $\{E_\alpha\}$ 是 Z 的子集，证明 $(\bigcup E_\alpha)_x=\bigcup(E_\alpha)_x$.

10. $\Omega=N=\{1,2,\cdots\}$，μ 是计点测度，$\mathscr{F}$ 是 N 的一切子集，$(\Omega,\mathscr{F},\mu)$ 是测度空间。设 $(Y,\mathscr{G},\nu)$ 是任意测度空间，证明 $Z=\Omega\times Y$ 中的子集 E，当且仅当 E 的每一个截口 $E_x\in\mathscr{G}$ 时，$E\in\mathscr{F}\times\mathscr{G}$. 还证明在这种情形下，存在唯一的乘积测度 π，且

$$\pi(E)=\sum_{n=1}^{\infty}\nu(E_n),\ \forall E\in\mathscr{F}\times\mathscr{G}$$

f 是 Z 到 R 的可测函数的充分必要条件是截口 f_x 均 $\mathscr{G}$ 可测。更进一步证明，当且仅当

$$\sum_{n=1}^{\infty}\int_Y|f_n|d\nu$$

收敛时 f 关于 π 可积，此时

$$\int_Z fd\pi=\sum_{n=1}^{\infty}\left[\int_Y f_n d\nu\right]=\int_Y\left[\sum_{n=1}^{\infty}f_n\right]d\nu$$

11. 设 X 和 Y 都是单位区间 $[0,1]$，$\mathscr{A}$ 与 $\mathscr{G}$ 都是 $[0,1]$ 中 Borel 子集所成的类。μ 是 $\mathscr{A}$ 上的 Lebesgue 测度，ν 是 $\mathscr{G}$ 上的计点测度。若 $D=\{(x,y):\ x=y\}$，证明 D 是 $Z=X\times Y$ 的可测子集，但

$$\int\nu(D_x)d\mu(x)\neq\int\mu(D_y)d\nu(y)$$

（如果 μ 与 ν 不是 σ 有限，则引理 1.10.5 可能不成立）

12. 若 F 是 D 的示性函数，D 是习题 1.9,11 中所讨论的集。证明如果 μ 与 ν 不是 σ 有限，Tonelli 定理可能不成立。

13. 用习题 1.9,10 中的例子说明当 $(Y,\mathscr{G},\nu)$ 是任意测度空间，$(N,\mathscr{F},\mu)$ 是计数测度 μ 和 N 自然数集组成的测度空间，Tonelli 定理成立。

14. 若 $a_{mn}\geqslant 0,\forall m,n\in N$，证明

$$\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}a_{mn}=\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}a_{mn}(\leqslant+\infty)$$

15. 设

$$a_{mn}=\begin{cases}+1, & \text{当 } m=n\\ -1, & \text{当 } m=n+1\\ 0, & \text{其他场合}\end{cases}$$

证明

$$\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}a_{mn}=0,\ \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}a_{mn}=1$$

（Fubini 定理中可积性的假设是不可少的）

16. 设 f 是 $(\Omega_1,\mathscr{F}_1,\mu)$ 上的可积函数，g 是 $(\Omega_2,\mathscr{F}_2,\nu)$ 上的可积函数，而且设

$$h(\omega_1,\omega_2)=f(\omega_1)g(\omega_2)$$

若 π 是 μ 与 ν 的乘积，证明 h 是 π 可积的，且

$$\int_{\Omega}hd\pi=\left[\int_{\Omega_1}fd\mu\right]\left[\int_{\Omega_2}gd\nu\right]$$

其中 $\Omega=\Omega_1\times\Omega_2$。

17. 假设 $(\Omega_1,\mathscr{F}_1,\mu)$ 与 $(\Omega_2,\mathscr{F}_2,\nu)$ 是 σ 有限的，E，F 都属于 $\mathscr{F}_1\times\mathscr{F}_2$。若

$$\nu(E_{\omega_1})=\nu(F_{\omega_1}),\ \forall\omega_1\in\Omega_1$$

证明 $\pi(E)=\pi(F)$。

第一章 测 验 题

1. 选择题

选择最符合题意的答案。

(1) 以原点为中心，$\frac{1}{n}$ 为半径的开圆记作 A_n，那么 $\limsup A_n=$

________。

A. 空集；B. 开单位圆；C. 闭单位圆；D. 非以上答案。

(2) 若 μ 是代数上非负有限可加集函数，$\sum_{n=1}^{\infty}A_n\in\mathscr{F}$（代数），则

________成立。

A. $\mu\left(\sum_{n=1}^{\infty}A_n\right)\leqslant\sum_{n=1}^{\infty}\mu(A_n)$；B. $\mu\left(\sum_{n=1}^{\infty}A_n\right)=\sum_{n=1}^{\infty}\mu(A_n)$；

C. $\mu\left(\sum_{n=1}^{\infty} A_n\right) \geqslant \sum_{n=1}^{\infty} \mu(A_n)$； D. $\mu\left(\sum_{n=1}^{\infty} A_n\right) \neq \sum_{n=1}^{\infty} \mu(A_n)$.

(3) 若 $A \subset \Omega$，$\chi_A(\omega)$ 是 A 的示性函数，则 χ_A________.

A. 可测； B. 不可测； C. 不一定可测.

(4) 若 $(\Omega, \mathscr{F})$ 为可测空间，f：$\Omega \to R$ 的函数，对任何 $a<b$，$\{\omega: a<f(\omega)\leqslant b\} \in \mathscr{F}$，则 f 关于 $\mathscr{F}$________.

A. 可测； B. 不可测； C. 不一定可测.

(5) 若 f 是 Ω 上的 $\mathscr{F}$ 可测函数，则 f^2 也是 $\mathscr{F}$ 可测函数，反之_____.

A. 亦然； B. 不然； C. f 不可测.

(6) 对任何实数 a，$\{\omega: f(\omega)=a\} \in \mathscr{F}$，则 f 是________.

A. 可测函数； B. 不可测函数； C. 不一定可测.

(7) f 可积，那么________.

A. f^2 可积； B. f^2 与 $|f|$ 均可积；

C. f^2 可积，$|f|$ 不可积； D. 非以上答案.

(8) 若 A_1 和 A_2 是广义测度 λ 的两个正集，那么 $A_1 \cup A_2$________.

A. 不是正集； B. 是正集；

C. 可能是负集； D. 不一定是正集.

(9) μ 是代数 $\mathscr{A}$ 上的测度，μ^* 是外测度，那么 μ^* 是________.

A. $\mathscr{M}(\mathscr{A})$ 上的外测度； B. $\mathscr{M}(\mathscr{A})$ 上的测度；C. $\mathscr{A}$ 上的外测度.

(10) 几乎处处收敛的函数列必依测度收敛. 这一命题________.

A. 正确； B. 不正确； C. 有条件地成立.

2. 证明题

(1) 设 $\mathscr{F}$ 是代数，μ 是 $\mathscr{F}$ 上的 σ 有限测度，

$$\mathscr{A}=\left\{A: A \in \mathscr{S}(\mathscr{F}) \text{ 且 } A_n \in \mathscr{F},\ \mu(A_n)<\infty, A \subset \bigcup_{n=1}^{\infty} A_n\right\}$$

这里 $\mathscr{S}(\mathscr{F})$ 表示包含 $\mathscr{F}$ 的最小 σ 代数，证明 $\mathscr{A}=\mathscr{S}(\mathscr{F})$.

(2) 设 $(\Omega, \mathscr{F}, \mu)$ 是有限测度空间，若 $f_n \xrightarrow[n\to\infty]{\mu} f$，证明 $f_n^2 \xrightarrow[n\to\infty]{\mu} f$，这里 f_n 和 f 均为几乎处处有限的非负可测函数.

(3) 若 $n\to\infty$，$\|f-f_n\|_p \to 0$，证明 $f_n \xrightarrow[n\to\infty]{\mu} f$.

第二章　鞅与鞅型序列

平稳随机过程、马尔可夫过程、鞅和随机微分方程理论都是近代控制理论的重要数学基础. 现有的随机过程书籍和随机系统方面的书籍都十分详细地讨论了平稳过程、马尔可夫过程理论尤其是平稳过程以及它们在控制理论中的应用，但对鞅和随机微分方程的内容却只作了简单的介绍. 有的随机系统书籍甚至于忽略了这一重要内容. 然而在近代控制理论文献中鞅和随机微分方程的内容愈来愈多地出现，并显示出了它们强有力的作用. 这样，编写以控制论为背景的鞅和随机微分方程书籍就显得十分必要. 本章和下面的两章，就是为解决这一客观需要而撰写的.

例如，在系统辨识、自适应控制、估计理论和信号处理中常常要考虑收敛性和渐近性态问题. 这些问题涉及到相依随机序列(或过程)的极限和渐近性态问题. 鞅与鞅型序列具有优良的极限性质和渐近性态，因此很自然地成了解决这些问题的基本工具之一.

在本书中，我们利用第一章的测度论知识，来讲解鞅的有关基本定理，这比阅读专著方便得多. 如果你在控制论的书籍、文章中遇到了超出本书范围的内容，由于你阅读了本书，有了一定的基础知识，因此在阅读参考书时会感到很方便.

本章首先介绍概率基础概念和条件数学期望，然后介绍鞅、停时、平方可积鞅等基本概念，最后证明鞅的基本收敛定理以及鞅型序列的有关收敛定理.

2.1　概率基础概念

摘要　本节旨在建立概率论中的基本概念与测度论中概念之间的对应关系. 概率是正则测度，随机变量是$\mathscr{F}$-可测函数，数学期望是积分等等. 同时，还引进了相互独立与无限维乘积空间的概念.

概率论是研究随机现象的数学理论. 所谓随机现象是指在一次试验前不能确定其结果，在大量试验中有明显规律性的那一类现象. 例如掷一硬币，其结果是国徽朝上(记作 H)还是字朝上(记作 W)，在投掷前是无法确定的. 但大量的试验表明：在大数次投掷下出现H和出现 W 的次数几乎相等. 人口统计表明:生男孩的概率与生女孩的概率大致相等，更精确地说男孩出生的概率为 0.517，而女孩出生的概率为 0.483.

在掷硬币试验中我们把出现H和出现W的情况都称为随机事件. 必然事件是出现H或 W，不可能事件为既不出现H又不出现 W. 这样我们把随机事件与可测空间联系了起来. 我们把

$$\Omega = \{H, W\}$$

称为必然事件，Ω 中的点 H，W 称为基本事件，而

$$\mathscr{F} = \{\Omega, H, W, \varnothing\}$$

组成σ代数. $\mathscr{F}$上的测度 P，具有 $P(\Omega) = 1$ 的性质，称为概率. 在这里

$$P(H) = P(W) = \frac{1}{2},\ P(\Omega) = 1,\ P(\varnothing) = 0$$

一般地说，$(\Omega, \mathscr{F}, P)$ 是一测度空间，若Ω中的点记作

ω，称为基本事件． 则$\mathscr{F}$是Ω的子集所成的σ代数．若$E \in \mathscr{F}$，那么E称为随机事件．$P(\Omega)=1$，P称为概率测度．

由于$P(\Omega)=1$，P也称为正则测度．概率测度P具有如下性质：

(1) $P(A) \geqslant 0, \ \forall A \in \mathscr{F}$；

(2) $P(\Omega)=1, \ P(\varnothing)=0$；

(3) 若 $A_i \in \mathscr{F}, \ i=1,2,\cdots, \ A_i$ 是两两不相交的集列，即 $A_i \cap A_j=\varnothing, \ \forall i \neq j$，那么

$$P\left(\sum_{i=1}^{\infty} A_i\right)=\sum_{i=1}^{\infty} P(A_i)$$

很明显这三条性质，也可以用以定义概率测度，即$\mathscr{F}$是σ代数，P是$\mathscr{F}$上的非负集函数，若满足以上三条则称为概率测度．

为了今后讨论的需要，我们把$\mathscr{F}$中任何概率为零的集合A的子集（$B \subset A$）B都算作随机事件，即若$P(A)=0$，$B \subset A$，则$B \in \mathscr{F}$且$P(B)=0$．这就是说，$(\Omega, \mathscr{F}, P)$是完备测度空间．

设 $A_i \in \mathscr{F}, i \geqslant 1$，如果对任意足标 $\{i_1, \cdots i_k\}$ 有

$$P\left(\bigcap_{j=1}^{k} A_{i_j}\right)=\prod_{j=1}^{k} P(A_{i_j}) \tag{2.1.1}$$

则称随机事件$A_1, A_2, \cdots$相互独立．

若$\mathscr{F}_1$是Ω中子集所成的σ代数，只要 $A \in \mathscr{F}_1$ 必有$A \in \mathscr{F}$，那么$\mathscr{F}_1$是$\mathscr{F}$的子σ代数．设$\mathscr{F}_i, i=1,2,\cdots$都是$\mathscr{F}$的子$\sigma$代数，如果对任意足标集$\{i_1, \cdots, i_k\}$以及任意$A_k \in \mathscr{F}_{i_k}$，$A_1, \cdots, A_k$都相互独立，那么称$\sigma$代数$\mathscr{F}_1$，$\mathscr{F}_2, \cdots$相互独立．

引理 2.1.1（Borel-Cantelli 引理） 设 $E_1, E_2, \cdots$ 表示一事件序列，如果

$$\sum_{i=1}^{\infty} P(E_i) < \infty \tag{2.1.2}$$

那么

$$P(\limsup_n E_n) = 0 \tag{2.1.3}$$

证明 因为

$$\limsup_n E_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} E_k$$

而 $\bigcup_{k=n}^{\infty} E_k$ 是一个下降序列，由测度的连续性[引理 1.3.2(2)]可知

$$P(\limsup_n E_n) = P\left(\lim_{n\to\infty} \bigcup_{k=n}^{\infty} E_k\right) = \lim_n P\left(\bigcup_{k=n}^{\infty} E_k\right)$$

$$\leqslant \lim_n \sum_{k=n}^{\infty} P(E_k) = 0 \text{ [由式 (2.1.2)]} \quad ∎$$

引理 2.1.2（**Borel-Cantelli 引理**） 设 $E_1, E_2, \cdots$ 是独立的事件序列，且

$$\sum_{n=1}^{\infty} P(E_n) = \infty \tag{2.1.4}$$

那么

$$P(\limsup_n E_n) = 1 \tag{2.1.5}$$

证明 因为

$$P(\limsup_n E_n) = \lim_{n\to\infty} P\left(\bigcup_{k=n}^{\infty} E_k\right)$$

$$= \lim_{n\to\infty} \left[1 - P\left(\bigcap_{k=n}^{\infty} E_k'\right)\right]$$

这里 $E_k' = \Omega - E_k$。

由事件独立性知

$$P\left(\bigcap_{k=n}^{\infty} E_k'\right)=\prod_{k=n}^{\infty}(1-P(E_k))$$

因为 $0 \leqslant P(E_k) \leqslant 1$，由不等式 $1-x \leqslant e^{-x}$，有

$$\prod_{k=n}^{\infty}(1-P(E_k)) \leqslant \prod_{k=n}^{\infty} e^{-P(E_k)}$$

$$=\exp\left(-\sum_{k=n}^{\infty} P(E_k)\right)$$

即

$$P\left(\bigcap_{k=n}^{\infty} E_k'\right) \leqslant \exp\left(-\sum_{k=n}^{\infty} P(E_k)\right)$$

因为

$$\sum_{k=n}^{\infty} P(E_k)=\infty$$

所以

$$P\left(\bigcap_{k=n}^{\infty} E_k'\right)=0$$

故

$$P(\limsup_n E_n)=1$$ ∎

推论 2.1.1 若 E_n 是独立事件序列，那么

$$P(\limsup_n E_n)=\begin{cases}0, & 当 \sum_{n=1}^{\infty} P(E_n)<\infty \\ 1, & 当 \sum_{n=1}^{\infty} P(E_n)=\infty\end{cases}$$

证明 由引理 2.1.1 和 2.1.2 自明. ∎

推论 2.1.1 称为 Borel 的零-壹律.

在直观概率论中，如果每次试验的结果，可以用一个实数 X 来表示，而且对于任何实数 x，"$X \leqslant x$" 有确定概率，则我

们称 X 是随机变量. 下面我们用测度论的术语进行分析:

首先注意定义中"每次试验的结果,可以用一个实数 X 来表示",这就是说,我们在基本空间 Ω 上定义了一个实值函数

$$X(\omega),\ \forall\omega\in\Omega$$

另外,定义中"$X\leqslant x$"有确定概率,是指对于一些最简单的复合事件,如

$$\{\omega\mid X(\omega)<x\}\tag{2.1.6}$$

它应该属于事件 σ 代数 $\mathscr{F}$,因而可以确定它的概率. 注意到(2.1.6)式和第一章的引理 1.2.1,不难发现用测度论的术语来说,随机变量就是关于 σ 代数 $\mathscr{F}$ 可测的可测函数.

随机变量的数学期望与方差是重要的数学特征. 数学期望指随机变量取值的平均数. 当随机变量只取有限多个值 $x_i(i=1,2,\cdots,n)$, 而取 x_i 值的概率为 p_i 时, $X(\omega)$ 的平均值应该是[用 $EX(\omega)$ 表示 $X(\omega)$ 的数学期望].

$$EX(\omega)=\sum_{i=1}^{n}x_ip_i\tag{2.1.7}$$

换一种写法: 设

$$X(\omega)=\sum_{i=1}^{n}x_i\chi_{A_i}$$

这里 $A_i\cap A_j=\varnothing(i\neq j)$, $\bigcup_{i=1}^{n}A_i=\Omega$. 于是 $X(\omega)$ 就是简单函数, 式(2.1.7)恰好是简单函数 $X(\omega)$ 关于测度 P 的积分. 类似地,当 $X(\omega)$ 取值于$(-\infty,+\infty)$时,类似于可测函数的积分应该有

$$EX(\omega)=\int_{\Omega}XdP\tag{2.1.8}$$

方差表示随机变量偏离其数学期望的程度, 可以用下式来定义:

$$\mathrm{Var}\, X(\omega) = E[X(\omega) - EX(\omega)]^2 \qquad (2.1.9)$$

我们把 $|X(\omega)|^r, r > 0$ 的数学期望称为随机变量 $X(\omega)$ 的 r 阶矩；$|X(\omega) - EX(\omega)|^r$ 的数学期望称为 r 阶的中心矩；$(X(\omega) - EX(\omega))(Y(\omega) - EY(\omega))$ 的数学期望称为随机变量 $X(\omega)$ 和 $Y(\omega)$ 的协方差，记为

$$\mathrm{Cov}(X(\omega)Y(\omega)) = E(X - EX(\omega))(Y(\omega) - EY(\omega))$$

假设 X, Y 是足够高次可积的随机变量，那么我们可以证明下列不等式：

引理 2.1.3（C_r 不等式）

$$E|X + Y|^r \leqslant C_r(E|X|^r + E|Y|^r) \qquad (2.1.10)$$

当 $r < 1$ 时 $C_r = 1$；当 $r \geqslant 1$ 时 $C_r = 2^{r-1}$。特别当 $r = 2$ 时有

$$E|X + Y|^2 \leqslant 2E|X|^2 + 2E|Y|^2 \qquad (2.1.11)$$

证明

(1) 先证明对 $a, b \in R$ 有

$$|a + b|^r \leqslant C_r|a|^r + C_r|b|^r$$

其中

$$C_r = \begin{cases} 1, & \text{当 } 0 < r \leqslant 1 \\ 2^{r-1}, & \text{当 } r \geqslant 1 \end{cases}$$

对 $r \geqslant 1$：作辅助函数 $\varphi(x) = x^r + (1-x)^r,\ x \in (0, 1)$，于是

$$\varphi'(x) = rx^{r-1} - r(1 - x)^{r-1}$$

在 $(0,1)$区间内，$\varphi(x)$ 有唯一的平衡点 $x = \frac{1}{2}$，即

$$\varphi'\left(\frac{1}{2}\right) = 0$$

容易验证 $$\varphi''\left(\frac{1}{2}\right) \geqslant 0$$

因此 $\varphi(x)$ 在 $(0,1)$ 内，$x=\dfrac{1}{2}$ 处取得最小值，即在 $x\in(0,1)$ 时

$$\varphi(x)\geqslant\varphi\left(\frac{1}{2}\right)=\frac{1}{2^{r-1}} \tag{2.1.12}$$

再令 $x=\dfrac{|a|}{|a|+|b|}$，由式 (2.1.12) 可知

$$\frac{1}{2^{r-1}}\leqslant\frac{|a|^r}{(|a|+|b|)^r}+\left(1-\frac{|a|}{|a|+|b|}\right)^r$$

于是

$$|a+b|^r\leqslant 2^{r-1}|a|^r+2^{r-1}|b|^r \tag{2.1.13a}$$

对 $r<1$：由于

$$\left(1+\left|\frac{b}{a}\right|\right)^{r-1}\leqslant 1,\ \left(1+\left|\frac{a}{b}\right|\right)^{r-1}\leqslant 1$$

因此

$$\begin{aligned}|a+b|^r&\leqslant(|a|+|b|)^r\\&=|a|^r\left(1+\left|\frac{b}{a}\right|\right)^{r-1}+|b|^r\left(1+\left|\frac{a}{b}\right|\right)^{r-1}\\&\leqslant|a|^r+|b|^r\end{aligned} \tag{2.1.13b}$$

由式 (2.1.13a) 和 (2.1.13b) 可得

$$|X+Y|^r\leqslant C_r|X|^r+C_r|Y|^r$$

由积分性质引理 1.4.2(a) 和推论 1.4.1 可知

$$E|X+Y|^r\leqslant C_rE|X|^r+C_r|Y|^r \qquad ▌$$

引理 2.1.4（基本不等式） 设 $r,a>0$，$E|X|^r$ 存在，那么

$$\frac{E|X|^r-a^r}{\sup|X|^r}\leqslant P\{|X|\geqslant a\}\leqslant\frac{E|X|^r}{a^r} \tag{2.1.14}$$

证明 令 $A=\{|X|\geqslant a\}$，设 $P(A)>0$

而且

$$E|X|^r = \int_A |X|^r dP + \int_{A^c} |X|^r dP$$

因为

$$0 < a^r P(A) \leqslant \int_A |X|^r dP \leqslant \sup|X|^r P(A) \quad (2.1.15)$$

所以

$$P(A) \leqslant \frac{\int_A |X|^r dP}{a^r} \leqslant \frac{\int |X|^r dP}{a^r} = \frac{E|X|^r}{a^r}$$

$$0 \leqslant \int_{A^c} |X|^r dP \leqslant a^r \quad (2.1.16)$$

由式 (2.1.15) 得

$$P(A) \geqslant \frac{\int_A |X|^r dP}{\sup|X|^r}$$

由式 (2.1.16) 得

$$\int_A |X|^r dP = E|X|^r - \int_{A^c} |X|^r dP \geqslant E|X|^r - a^r$$

因此

$$P(A) \geqslant \frac{E|X|^r - a^r}{\sup|X|^r}$$ ▌

习　题　2.1

1.设 $(\Omega, \mathscr{F}, P)$ 为概率空间，$\{X_n\}$ 是随机变量序列，X 是可积随机变量，若

$$\int |X_n - X|^r dP \to 0, \ \forall r > 0$$

则称 X_n 为 L_r 收敛于 X，记作 $X_n \xrightarrow{L_r} X$.

证明 (1) 若 $X_n \in L_r$，$E|X_n - X|^r \to 0$，那么 $X \in L_r$；

(2) 若 $X_n \xrightarrow{L_r} X$，$0 < r' \leqslant r$，那么 $X_n \xrightarrow{L_{r'}} X$.

2. 用图总结几乎必然收敛，依测度收敛与 L_r 收敛之间的关系.

3. 把测度论术语与概率术语之间的关系写完全.

4. 若 g_ε 是 $[0,\infty)$ 上的非负 Borel 可测函数，对于 $x\geqslant\varepsilon$, $g_\varepsilon(x)\geqslant g_\varepsilon(\varepsilon)$，证明

$$P[\,|X|\geqslant\varepsilon\,]\leqslant\frac{Eg_\varepsilon(|X|)}{g_\varepsilon(\varepsilon)}$$

5. 证明，当且仅当 $|X|$ 是退化的随机变量(即几乎必然为常数) 时，关于 r 的凸函数 $\log E|X|^r$ 是 r 的线性函数.

6. 李亚普诺夫不等式： 设 $\mu_r=E|X|^r$，若 $r\geqslant s\geqslant t\geqslant 0$，证明 $\mu_t^{r-s}\mu_r^{s-t}\mu_s^{t-r}\geqslant 1$.

7. 证明：当且仅当存在一个 ε_n 序列，且 $\varepsilon_n\to 0$ 使得

$$P\Big(\bigcup_{k\geqslant n}[\,|X_k-X|\geqslant\varepsilon_k\,]\Big)\to 0\text{时}X_n\xrightarrow{a.s.}X$$

2.2 分布函数与特征函数

提要 本节给出概率分布函数的定义以及随机变量序列依分布律收敛 (或简称分布收敛)的概念. 如果随机变量序列依测度收敛，那么它一定依分布律收敛. 反之，不一定成立. 本节还引入了随机变量特征函数的概念，并证明了分布函数与特征函数间的一一对应关系.

分布函数是描述随机变量的 R 上的点函数.

定义 2.2.1 在 R 上单调非降的右连续函数 $F(x)$，如果满足

$$\lim_{x\to-\infty}F(x)=0,\quad \lim_{x\to\infty}F(x)=1 \qquad (2.2.1)$$

则 $F(x)$ 称为概率分布函数.

概率分布函数是一种定分布函数，在概率论中，通常简称为分布函数.

假定 $X(\omega)$ 是概率空间 $(\Omega,\mathscr{F},P)$ 上的随机变量. 我们定义一个实函数

$$F(x)=P(X(\omega)\leqslant x) \tag{2.2.2}$$

很显然,它是一非降函数,且

$$\lim_{x\to-\infty}F(x)=P(\varnothing)=0,\quad \lim_{x\to\infty}F(x)=P(\Omega)=1$$

另外,由式(2.2.2)知,当 $\Delta x\to 0$ 时

$$F(x+\Delta x)-F(x)=P(x<X(\omega)\leqslant x+\Delta x)\to 0$$

因此由式(2.2.2)定义的函数 $F(x)$ 是右连续的,而且是分布函数.

这说明概率空间 $(\Omega,\mathscr{F},P)$ 上任一随机变量都能确定一个分布函数. 确定了分布函数之后,我们令

$$P_x((a,b])=F(b)-F(a) \tag{2.2.3}$$

不难证明式(2.2.3)在 Borel 代数上定义了一个测度(利用测度扩张定理),同时它也给出了 R 上的概率空间 $(R,\mathscr{B},P_x)$, 这里 $\mathscr{B}$ 表示 R 上的 Borel 代数. 反过来,若给出了分布函数 $F(x)$,就存在一个概率空间 $(R,\mathscr{B},\mu)$ 及其上的随机变量

$$X=X(\omega),\ \text{值域为}\ R$$

使得

$$F(x)=P(X(\omega)\leqslant x),\quad x\in R$$

上面的讨论说明: 给出了概率空间 $(\Omega,\mathscr{F},P)$ 上的随机变量 $X(\omega)$, 就确定了 $(R,\mathscr{B})$ 上的一个测度. 因此也确定了一个概率空间 $(R,\mathscr{B},P_x)$, 而且下面的定理成立.

定理 2.2.1 如果 $F(x)$ 是由随机变量 $X(\omega)$ 确定的分布函数, $g(x)$ 是 R 上任一有限 Borel 可测函数, 那么 $g(X(\omega))$ 也是随机变量,且

$$Eg(X(\omega))=\int_R g(s)dF(s) \tag{2.2.4}$$

证明 由于 g 是 Borel 可测函数，因而对任一 Borel 集 $E \in \mathscr{B}$，逆映象 $g^{-1}(E) \in \mathscr{B}$。于是

$$\{\omega | g(X(\omega)) \in E\} = \{\omega | X(\omega) \in g^{-1}(E)\} \in \mathscr{F}$$

所以 $g(X(\omega))$ 仍是随机变量。为了证明式 (2.2.4)，首先我们要指出式 (2.2.4) 的右端是 Lebesgue-Stieltjes 积分。对于这种积分单调收敛定理仍然成立。

先证明在 $g(x) = \chi_E$ 的情形下，式 (2.2.4) 成立，这里 $E \in \mathscr{B}$。事实上，这时

$$Eg(X(\omega)) = \int_{\Omega} \chi_{\{\omega | X(\omega) \in E\}} dP = P\{\omega | X(\omega) \in E\}$$

因为

$$P_x(E) = P\{\omega | X(\omega) \in E\}$$

于是

$$\int_R g(x) dF(x) = \int_R \chi_E dP_x = P_x(E)$$

这就证明了式 (2.2.4) 对一切示性函数都成立。简单函数是有限个示性函数(不相交集的示性函数)的线性组合。因此，式 (2.2.4) 对简单函数也成立。根据非负 Borel 可测函数是某一简单函数列的极限，以及单调收敛定理，式 (2.2.4) 对于非负 Borel 可测函数成立。

又由于

$$g(x) = g^+(x) - g^-(x)$$

因此对于一切 Borel 可测函数 $g(x)$，式 (2.2.4) 也成立。

定义 2.2.2 当分布函数 $F(x)$ 可微时

$$f(x) = \frac{d}{dx} F(x) \tag{2.2.5}$$

$f(x)$ 称为概率分布密度或对应随机变量的分布密度。这时的积分表达式 (2.2.4) 可写成

$$Eg(X(\omega)) = \int_R g(x) f(x) dx \tag{2.2.6}$$

定义 2.2.3 设随机变量序列 $\{X_n(\omega)\}$ 对应的分布函数列为 $\{F_n(x)\}$，随机变量 $X(\omega)$ 的分布函数为 $F(x)$。如果在 $F(x)$ 的一切连续点上都有

$$\lim_{n\to\infty} F_n(x) = F(x) \tag{2.2.7}$$

则称 $X_n(\omega)$ 依分布律收敛于 $X(\omega)$，简记作

$$X_n(\omega) \xrightarrow[n\to\infty]{\text{d. f.}} X(\omega)$$

注意：由于 $F(x)$ 是有界非降函数，因而它的不连续点至多有可列个. 这是因为如令

$$dF = \lim_{\Delta x\to 0} F(x+\Delta x) - \lim_{\Delta x\to 0} F(x-\Delta x)$$

则 $dF > \frac{1}{2}$ 的是不可能多于 2 个. 同理，$dF > \frac{1}{n}$ 的点不可能多于 n 个. 令 $n\to\infty$，$F(x)$ 的所有不连续点至多是可列个.

这样，随机变量（$\mathscr{F}$ 可测函数）的收敛关系又多了一种. 那么分布收敛与其他收敛的关系怎样呢？

定理 2.2.2 如果随机变量序列 $\{X_n(\omega)\}$ 依测度收敛于 $X(\omega)$，即 $X_n \xrightarrow{\mu} X$，那么 $X_n \xrightarrow{\text{d.f.}} X$.

证明 由于

$\{\omega | X(\omega) \leqslant a\} = \{\omega | X_n(\omega) \leqslant b; X(\omega) \leqslant a\}$

$\cup \{\omega | X_n(\omega) > b; X(\omega) \leqslant a\}$

$\subset \{\omega | X_n(\omega) \leqslant b\} \cup \{\omega | X_n(\omega) > b; X(\omega) \leqslant a\}$

于是有

$$P\{\omega | X(\omega) \leqslant a\} \leqslant$$
$$P\{\omega | X_n(\omega) > b; X(\omega) \leqslant a\} + F_n(b) \tag{2.2.8}$$

设 $a < b$，由 $X_n \xrightarrow{P} X$

$$P\{\omega \mid X_n(\omega) > b; X(\omega) \leqslant a\} \leqslant P\{\omega \mid |X_n - X| \geqslant b - a\} \xrightarrow[n\to\infty]{} 0$$

在式(2.2.8)两端取下极限得

$$F(a) \leqslant \liminf_{n\to\infty} F_n(b)$$

用同样的方法可以得到

$$\limsup F_n(b) \leqslant F(c), \quad b < c$$

对于 $a < b < c$ 有

$$F(a) \leqslant \liminf_{n\to\infty} F_n(b) \leqslant \limsup_{n\to\infty} F_n(b) \leqslant F(c)$$

如果 b 是 $F(x)$ 的连续点，令 $a \uparrow b$ 及 $c \downarrow b$，那么

$$F(b) = \lim_{n\to\infty} F_n(b)$$

这就是说

$$X_n \xrightarrow[n\to\infty]{\text{d.f.}} X \qquad \blacksquare$$

定理 2.2.3 设 $X_n \xrightarrow{\text{d.f.}} X$，$X_n$ 和 X 的分布函数分别为 $F_n(x)$ 和 $F(x)$，g 是 R 上的有界连续函数，那么

$$\int_R g(x)dF_n(x) \to \int_R g(x)dF(x) \tag{2.2.9}$$

证明 首先在闭区间$[a,b]$上作一分割

$$a \leqslant x_{m_1} < x_{m_2} < \cdots < x_{m_{k+1}} = b$$

当 $m_k \to \infty$ 时，$\Delta = \sup_k (x_{m_k} - x_{m_{k-1}}) \to 0$。设

$$g_m = \sum_{j=1}^{k(m)} g(x_{m_j}) \chi_{(x_{m_j}, x_{m_{j+1}}]}$$

根据 Riemann-Stieltjes 积分定义，当 $m \to \infty$ 时有

$$\int_a^b g_m dF_n \to \int_a^b g dF_n, \quad \int_a^b g_m dF \to \int_a^b g dF \tag{2.2.10}$$

我们把所有的分点选为 F 的连续点. 由于 $X_n \xrightarrow{\text{d.f.}} X$，因

而对一切 m_k，只要 $n\to\infty$ 就有

$$F_n(x_{m_k}) - F(x_{m_k}) \xrightarrow[n\to\infty]{} 0$$

因此

$$\int_a^b g_m dF_n = \sum_{j=1}^{k(m)} g(x_{m_j})[F_n(x_{m_{j+1}}) - F_n(x_{m_j})]$$

$$\xrightarrow[n\to\infty]{} \sum_{j=1}^{k(m)} g(x_{m_j})[F(x_{m_{j+1}}) - F(x_{m_j})] = \int_a^b g_m dF \tag{2.2.11}$$

$$\int_a^b g dF_n - \int_a^b g dF = \int_a^b (g - g_m) dF_n + \int_a^b g_m dF_n - \int_a^b g_m dF + \int_a^b (g_m - g) dF \tag{2.2.12}$$

对式 (2.2.10) 和 (2.2.11)，若令 $m\to\infty$，$n\to\infty$，式 (2.2.12) 右端就趋于零，即

$$\int_a^b g dF_n - \int_a^b g dF \xrightarrow[n\to\infty]{} 0$$

再考虑整个数直线 R 上的情形：因为

$$\left|\int g dF_n - \int g dF\right| \leqslant \left|\int g dF_n - \int_a^b g dF_n\right| + \left|\int_a^b g dF - \int_a^b g dF_n\right| + \left|\int_a^b g dF - \int g dF\right| \tag{2.2.13}$$

根据上面的证明，式(2.2.12)右端第二项当 $n\to\infty$ 时趋于零. 第一项满足

$$\left|\int g dF_n - \int_a^b g dF_n\right| \leqslant M(P_{x_n}(-\infty, a] + P_{x_n}(b, +\infty)) = M[F_n(a) + 1 - F_n(b)]$$

其中 M 是 g 的上界，当 $a\to-\infty$，$b\to\infty$ 时，$[F_n(a) + 1 - F_n(b)]\to 0$.

对于式 (2.2.13)，使 $|a|$，$|b|$ 足够大，取 $n\to\infty$，这样

就可以使式 (2.2.13) 右端任意小。 ▮

特征函数是研究随机变量序列极限，求解随机变量分布函数和矩的有力工具。

定义 2.2.4 对任意分布函数 F，函数

$$\phi(u)=\int e^{iux}dF(x)=\int\cos ux\,dF(x)+i\int\sin ux\,dF(x) \tag{2.2.14}$$

称为 F 的特征函数。这里 i 表示虚单位，即 $i^2=-1$。当 $F(x)$ 是随机变量 $X(\omega)$ 的分布函数时，$\phi(u)$ 也称为随机变量 $X(\omega)$ 的特征函数。特征函数是一个实变量的复值函数。

由定理 2.2.1 可知

$$\phi(u)=\int e^{iux}dF(x)=Ee^{iuX(\omega)}$$

显然每一个分布函数可以对应一个特征函数。特征函数具有如下性质：

(1) 特征函数在$(-\infty,+\infty)$上一致连续，且有

$$\phi(0)=1,\ |\phi(u)|\leqslant 1,\ \forall u\in(-\infty,\infty) \tag{2.2.15}$$

证明 由于

$$e^0=1,\quad |e^{iux}|\leqslant 1$$

$$\int dF(x)=F(+\infty)-F(-\infty)=1$$

因此很容易由 $\phi(u)$ 的定义证明式 (2.2.15)。往证。$\phi(u)$的一致连续性。

因为

$$|\phi(u+h)-\phi(u)|=\left|\int e^{iux}(e^{ihx}-1)dF(x)\right|$$

$$\leqslant\int|e^{ihx}-1|dF(x)$$

对任意给定的 $\varepsilon>0$，可选取 $a>0$ 使得

$$\int_{|x|\geq a} dF(x) < \frac{\varepsilon}{4}$$

然后再取 $\delta > 0$，使得当 $|h| < \delta$ 时，

$$|e^{ihx} - 1| = |e^{i\frac{hx}{2}}(e^{i\frac{hx}{2}} - e^{-i\frac{hx}{2}})| = \left|2\sin\frac{hx}{2}\right| < \frac{\varepsilon}{2}$$

因此对任给的 $\varepsilon > 0$，可找到 $\delta>0$，和 $a > 0$，当 $|h| < \delta$ 时，

$$|\phi(u+h) - \phi(u)| \leqslant \int_{-a}^{a} |e^{ihx} - 1| dF + 2\int_{|x|\geq a} dF(x)$$

$$< \frac{\varepsilon}{2}[F(a) - F(-a)] + \frac{\varepsilon}{2} < \varepsilon$$

由于 δ 与 u 无关，因此 $\phi(u)$ 一致连续。 ▌

(2) 设 $\eta = a\xi + b$，a，b 为常数，那么

$$\phi_\eta(u) = e^{ibu}\phi_\xi(au) \tag{2.2.16}$$

其中 $\phi_\eta(u)$ 和 $\phi_\xi(u)$ 分别表示随机变量 η 和 ξ 的特征函数。

证明 根据特征函数的定义，有

$$\phi_\eta(u) = Ee^{iu\eta} = Ee^{iu(a\xi+b)} = e^{ibu}Ee^{iua\xi} = e^{ibu}\phi_\xi(au) \quad ▌$$

(3) 两个相互独立的随机变量之和的特征函数等于各个特征函数之积。

证明 设 ξ 和 η 是两个相互独立的随机变量，那么 $e^{iu\xi}$ 和 $e^{iu\eta}$ 也是独立的随机变量。

设 $\zeta = \eta + \xi$，于是

$$Ee^{iu\zeta} = Ee^{iu\xi} \cdot e^{iu\eta} = Ee^{iu\xi} \cdot Ee^{iu\eta}$$

因为独立随机变量乘积的数学期望等于它们的数学期望之乘积，因此

$$\phi_\zeta(u) = \phi_\xi(u) \cdot \phi_\eta(u) \quad ▌$$

(4) 设随机变量 ξ 有 n 阶原点矩，于是它的特征函数 $\phi(u)$ n 次可微，且当 $k \leqslant n$ 时

$$\phi^{(k)}(0) = i^k E\xi^k$$

证明 设

$$\phi(u)=\int e^{iux}dF(x)$$

因为

$$\left|\frac{\partial^k}{\partial u^k}e^{iux}\right|=|x^k i^k e^{-iux}|\leqslant |x|^k$$

根据假设条件

$$\int |x|^k dF(x)<\infty$$

由推论 1.5.4 知

$$\phi^{(k)}(u)=i^k\int x^k e^{iux}dF(x)$$

这里 $\phi^{(k)}(u)$ 表示 $\phi(u)$ 的 k 阶导数，于是

$$\phi^{(k)}(0)=i^k E\xi^k \qquad ∎$$

利用 (4)，人们很容易从特征函数求得随机变量的 k 阶原点矩。

引理 2.2.1 设 ξ 是 $N(0,1)$ 正态随机变量，即 ξ 的概率密度为

$$f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

那么它的特征函数为 $\phi(t)=e^{-t^2/2}$

证明 根据数学分析中 Gamma 函数的定义

$$\Gamma(s)=\int_0^\infty x^{s-1}e^{-x}dx,\quad s>0$$

以及 Gamma 函数的基本性质

$$\Gamma(s+1)=s\Gamma(s),\ \Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}$$

我们先考察下述积分：

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}x^{2n}e^{-x^2/2}dx$$

作变量代换，令 $z=\frac{x^2}{2}$，$dz=xdx$，于是

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}x^{2n}e^{-\frac{x^2}{2}}dx=\frac{2}{\sqrt{2\pi}}\int_0^{\infty}e^{-\frac{x^2}{2}}x^{2n}dx$$

$$=\frac{2}{\sqrt{2\pi}}\int_0^{\infty}(2z)^{n-\frac{1}{2}}e^{-z}dz$$

$$=\frac{2^n}{\sqrt{\pi}}\int_0^{\infty}z^{n-\frac{1}{2}}e^{-z}dz=\frac{2^n}{\sqrt{\pi}}\Gamma\left(n+\frac{1}{2}\right)$$

$$=(2n-1)!!$$

现在就可以求特征函数：

$$\phi(t)=Ee^{it\xi}=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{itx}e^{-\frac{x^2}{2}}dx$$

$$=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\sum_{n=0}^{\infty}\frac{(itx)^n}{n!}e^{-\frac{x^2}{2}}dx$$

$$=\sum_{n=0}^{\infty}\frac{(it)^n}{n!}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}x^ne^{-x^2/2}dx$$

$$=\sum_{n=0}^{\infty}\frac{(it)^{2n}}{(2n)!}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}x^{2n}e^{-x^2/2}dx$$

$$=\sum_{n=0}^{\infty}\frac{(it)^{2n}}{(2n)!}(2n-1)!!=\sum_{n=0}^{\infty}\frac{(it)^{2n}}{(2n)!}\frac{(2n)!}{2^nn!}$$

$$=\sum_{n=0}^{\infty}\left(-\frac{t^2}{2}\right)^n\cdot\frac{1}{n!}=e^{-t^2/2}\qquad(2.2.17)\blacksquare$$

例 1 设 ξ 是 $N(a,\sigma^2)$ 正态随机变量，其中 $a=E\xi$，$\sigma^2=E(\xi-a)^2$，求 ξ 的特征函数以及 $E|\xi-a|^{2n}$ 的表达式。

解 令 $\eta=\frac{\xi-a}{\sigma}$，$\eta$ 服从 $N(0,1)$ 正态分布。根据式(2.2.15)以及

$$\xi = \sigma\eta + a$$

可得

$$\phi_\xi(u) = e^{iau}\phi_\eta(\sigma u) = e^{iau} \cdot e^{-\frac{\sigma^2 u^2}{2}} = e^{iau-\frac{\sigma^2 u^2}{2}}$$

这里 $\phi_\xi(u)$，$\phi_\eta(u)$ 分别表示随机变量 ξ 和 η 的特征函数。令 $\zeta = \xi - a = \sigma\eta$，那么

$$\phi_\zeta(u) = e^{-\frac{\sigma^2 u^2}{2}}$$

于是求得

$$i^{2n}\phi^{(2n)}(0) = E|\xi - a|^{2n} = E\zeta^{2n} = (2n-1)!!\sigma^{2n}$$

例如

$$i^2\phi''(u) = \sigma^2 e^{-\frac{\sigma^2 u^2}{2}} - \sigma^4 u^2 e^{-\frac{\sigma^2 u^2}{2}}$$

所以

$$E|\xi - a|^2 = i^2\phi''(0) = \sigma^2$$ ∎

例 2 设 ξ 是 Poisson 随机变量

$$P(\xi = k) = \frac{e^{-\lambda}\lambda^k}{k!}, \quad k = 0, 1, 2, \cdots$$

于是

$$Ee^{it\xi} = \sum_{k=0}^{\infty} e^{itk}\frac{e^{-\lambda}\lambda^k}{k!} = e^{-\lambda}\sum_{k=0}^{\infty}\frac{(\lambda e^{it})^k}{k!}$$
$$= e^{-\lambda}\exp\{\lambda e^{it}\} = \exp\{\lambda(e^{it} - 1)\}$$

它是 Poisson 随机变量的特征函数。

从上面的例子看到，一个分布函数可以确定一个特征函数。但是，特征函数能否唯一决定分布函数呢？下面叙述的逆转公式和唯一性定理，就给出了这个问题以肯定答案。

定理 2.2.4（反演公式） 设随机变量 ξ 的特征函数和分布函数分别为 $\phi(u)$ 和 $F(x)$，又 a 和 b 是 $F(x)$ 的连续点，于是

(1) $F(b) - F(a) = \frac{1}{2\pi} \lim_{t\to\infty} \int_{-t}^{t} \frac{e^{ita} - e^{-itb}}{it} \phi(t)dt$

$$(2.2.18)$$

(2) 若 $\int_{-\infty}^{\infty} |\phi(t)| dt < \infty$，则分布函数 $F(x)$ 具有密度函数 $f(x)$，且

$$F(x) = \int_{-\infty}^{x} f(y)dy \qquad (2.2.19)$$

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx}\phi(t)dt \qquad (2.2.20)$$

证明 见一般概率论书籍(如文献[2] p. 369)。 ▮

定理 2.2.5 分布函数由其特征函数唯一决定。

证明 见文献[2]。 ▮

由此可见随机变量的分布函数与其特征函数是一一对应的。

下面我们给出常用分布函数与其特征函数的对应表

分布函数名称	特征函数 $\phi(t)$
退化分布 $\delta(x-a)$	e^{ita}
二项式分布 $C_n^k p^k q^{n-k}$	$(q+pe^{it})^n$
负二项式分布	$\{p[1-qe^{it}]^{-1}\}^r$
Poisson 分布	$\exp\{\lambda(e^{it}-1)\}$
$[-a,a]$ 上的均匀分布	$\frac{\sin at}{at}$
Laplace 分布	$\frac{1}{1+t^2}$
正态分布 $N(\mu,\sigma^2)$	$\exp\left[i\mu t - \frac{\sigma^2 t^2}{2}\right]$
Caudy 分布	$e^{-\|t\|}$
Gammar 分布	$\left(1-\frac{it}{\theta}\right)^{-\lambda}$
Beta 分布	$\frac{\Gamma(p+q)}{\Gamma(q)} \sum_{j=0}^{\infty} \frac{\Gamma(p+j)}{\Gamma(p+q+j)\Gamma(j+1)} (it)^j$

习 题 2.2

1.设 $\phi(u)$ 是特征函数，证明

$$\phi(-u)=\overline{\phi(u)}$$

这里 $\overline{\phi(u)}$ 表示 $\phi(u)$ 的共轭复函数.

2. Bochner-Khinchin 定理指出：满足 $\phi(0)=1$ 的连续函数为特征函数的充分必要条件是它是正定的.

一个连续函数 $\phi(t)$ 在 $-\infty<t<\infty$ 内正定是指对任何实数 $t_1, t_2,\cdots,t_n$，复数 $z_1,z_2,\cdots,z_n$ 以及整数 n，恒有

$$\sum_{k=1}^{n}\sum_{j=1}^{n} f(t_k-t_j)z_j\bar{z}_j\geqslant 0$$

(1) 如果 $\varphi_k(t)$，$k\geqslant 1$ 是特征函数，设 λ_k 为非负常数，且 $\sum\lambda_k=1$，证明 $\sum\lambda_k\phi_k(t)$ 是特征函数.

(2) 若 $\phi(t)$ 是特征函数，问 $\mathrm{Re}\phi(t)$ 和 $\mathrm{Im}\phi(t)$ 是否仍然为特征函数?

3.证明：当且仅当 $F_n(x)$ 分布收敛于

$$F(x)=\begin{cases}0, & x\leqslant C\\ 1, & x>C\end{cases}$$

时 $X_n\xrightarrow{p}C$.

4.设 $X_n=\begin{bmatrix}1 & 0\\ \frac{1}{2} & \frac{1}{2}\end{bmatrix}$，即 X_n 取 1 和 0 的概率各为 1/2，

$X=\begin{bmatrix}0 & 1\\ \frac{1}{2} & \frac{1}{2}\end{bmatrix}$，即 X 取 0 和 1 的概率为 1/2.

这里 $F_n(x)=F(x)$，但 $|X_n-X|=1$，这是哪一个命题的反例?

2.3 随机向量及多维正态随机变量

提要： 本节讨论随机向量、多维正态随机变量

的概念，并总结测度论概念与概率论概念间的联系。

设R表示数直线，$\mathscr{B}$表示 Borel 代数。 $R\times R$是一切有序点偶(x_1,x_2)（其中$x_1,x_2\in R$）的全体，称为二维欧氏空间，记作

$$R^2=R\times R=\{(x_1,x_2)|x_1\in R,x_2\in R\}$$

若$(R,\mathscr{B})$为 Borel 可测空间，由可测矩阵$(A\in\mathscr{B},B\in\mathscr{B})A\times B$生成的$\sigma$代数，记作$\mathscr{B}^2=\mathscr{B}\times\mathscr{B}$，$\mathscr{B}^2$中的集称为 Borel 可测集。

更一般地，设R^l表示l维欧氏空间，$\mathscr{B}^l$是R^l子集所成的 Borel σ 代数或简称为 Borel 代数。

$X=X(\omega)$是定义于Ω，取值于R^l的函数，如果对于任何$E\in\mathscr{B}^l$有

$$\{\omega|X(\omega)\in E\}\in\mathscr{F} \tag{2.3.1}$$

则称X为向量值可测函数。

如果$(\Omega,\mathscr{F},P)$是概率空间，对任何$E\in\mathscr{B}^l$式(2.3.1)成立，则称$X(\omega)$为l维随机向量。

设Y是l维随机向量，$\mathscr{F}_Y$表示包含一切形如

$$\{Y^{-1}(B)|B\in\mathscr{B}^l\} \tag{2.3.2}$$

集的最小σ代数，那么$\mathscr{F}_Y$称为Y的自然σ代数，或称为Y生成的σ代数。$\mathscr{F}_Y$记作$\sigma(Y)$或$\sigma(y_1,\cdots,y_l)$，其中y_i表示Y的第i个分量，$i=1,2,\cdots,l$。

对于随机向量$X:\Omega\to R^n$，若

$$P[X\in G]=\int_G p(x_1,\cdots,x_n)dx_1\cdots dx_n,\forall G\in\mathscr{B}^n$$

且

$$p(x_1,\cdots,x_n)=\frac{\sqrt{\det A}}{(2\pi)^{n/2}}\exp\left\{-\frac{1}{2}(X-m)^TA(X-m)\right\} \tag{2.3.3}$$

其中 $m=(m_1,\cdots,m_n)^T\in R^n$，det(·) 表示矩阵的行列式，$C^{-1}=A=[a_{jk}]\in R^{n\times n}$ 是正定矩阵，而且

$$EX=m \tag{2.3.4}$$

$$A^{-1}=C=[C_{jk}] \tag{2.3.5}$$

这里

$$C_{jk}=E[(X_j-m_j)(X_k-m_k)] \tag{2.3.6}$$

X_j 是 X 的第 j 个分量，则称 X 为随机正态向量。

定义 2.3.1 随机向量 $X:\Omega\to R^n$ 的特征函数 $\phi_X:R^n\to C$（C 表示复数域）定义为

$$\phi_X(u_1,\cdots,u_n)=E\{\exp[i(u_1X_1+\cdots+u_nX_n)]\}$$
$$=\int e^{i\langle u,x\rangle}dP_x \tag{2.3.7}$$

其中 $\langle u,x\rangle=u_1x_1+\cdots+u_nx_n$，当 X 为连续型随机向量时，式 (2.3.7) 表示为

$$\phi_X(u)=\int_{R_n}e^{i\langle u,x\rangle}p(x_1,\cdots,x_n)dx_1\cdots dx_n$$

当 X 为离散型随机向量时，式 (2.3.7) 表示为

$$\phi_X(u)=\sum_{k_1=1}^{\infty}\cdots\sum_{k_n=1}^{\infty}e^{i\langle u,x_k\rangle}p(X_1=x_{k_1}\cdots X_n=x_{k_n})$$

可以证明正态向量 $X=(X_1,\cdots,X_n)^T$，式 (2.3.3) 的特征函数为

$$\phi_X(u)=\exp\left\{\frac{1}{2}\sum_{j,k}c_{jk}u_ju_k+i\sum_j u_jm_j\right\} \tag{2.3.8}$$

由式 (2.3.3) 或 (2.3.8) 可知，正态随机向量（变量）完全由它的一阶和二阶矩所完全确定。

在研究条件数学期望和鞅论之前，我们来总结一下概率论中基本概念与测度论中概念的关系：

概率空间$\longleftrightarrow$正则完备测度空间

基本事件$\longleftrightarrow$空间Ω中的点

随机事件$\longleftrightarrow$可测集

必然事件$\longleftrightarrow$全空间Ω

不可能事件$\longleftrightarrow$空集

概率$\longleftrightarrow$正则测度（$P(\Omega)=1$）

几乎必然 a.s.依概率 1 成立，w.p.1$\longleftrightarrow$ 几乎处处 a.e.

随机变量$\longleftrightarrow\mathscr{F}$ 可测函数

数学期望$\longleftrightarrow$可测函数关于P的积分

依概率收敛$\longleftrightarrow$依测度收敛

分布收敛$\longleftrightarrow$在分布函数连续点收敛

测度论中的一些著名定理可改写如下：设 X_n，X，Y 都表示（$\Omega,\mathscr{F},P$）上的随机变量。

定理 2.3.1（单调收敛定理） （1）若 $X_n\geqslant Y,\forall n\geqslant 1$，$EY>-\infty$ 且 $X_n\uparrow X$（X_n 单调升地收敛到 X），那么

$$EX_n\xrightarrow[n\to\infty]{}EX$$

（2）若 $X_n\leqslant Y$，$n\geqslant 1$，$EY<\infty$ 且 $X_n\downarrow X$，那么

$$EX_n\xrightarrow[n\to\infty]{}EX$$

证明 （1）因为 $X_n\geqslant Y$，故 $X_n-Y\geqslant 0$，且$(X_n-Y)\uparrow(X-Y)$。

根据第一章单调收敛定理 1.4.1

$$E(X_n-Y)\xrightarrow[n\to\infty]{}E(X-Y)$$

于是得

$$EX_n\to EX(n\to\infty)$$

同理可证(2)。∎

注意：（1）定理 2.3.1 与第一章的定理 1.4.1 不同，因为在定理 2.3.1 中 X_n 不一定为非负可测函数。

(2) 定理 2.3.1 与第一章中的控制收敛定理也不同,因为在控制收敛定理中要求 $|X_n| < Y$,Y 为可积函数.

引理 2.3.1 (Fatou 引理) 设 $X_1, X_2, \cdots$ 是随机变量.

(1) 若 $X_n \geqslant Y$, $n \geqslant 1$ 且 $EY > -\infty$, 那么

$$E\lim\inf X_n \leqslant \lim\inf EX_n$$

(2) 若 $X_n \leqslant Y$, $n \geqslant 1$ 且 $EY < \infty$, 那么

$$\lim\sup EX_n \leqslant E\lim\sup X_n$$

证明

(1) 因为 $X_n \geqslant Y$, $X_n - Y \geqslant 0$, 根据引理 1.4.3, 可知

$$E\lim\inf(X_n - Y) \leqslant \lim\inf E(X_n - Y)$$

于是得

$$E\lim\inf X_n \leqslant \lim\inf EX_n$$

(2) 因为 $X_n \leqslant Y$, $Y - X_n \geqslant 0$, 与(1)类似可得

$$E\lim\inf(Y - X_n) \leqslant \lim\inf E(Y - X_n)$$

即

$$EY - E\lim\sup X_n \leqslant EY - \lim\sup EX_n$$

故

$$E\lim\sup X_n \geqslant \lim\sup EX_n \qquad \blacksquare$$

注意: 在引理 2.3.1 中并没有 $X_n \geqslant 0$ 的要求.

定理 2.3.2 (Lebesgue 控制收敛定理) 若 X_n 依概率收敛于 X, 即 $X_n \xrightarrow{P} X$ 且 $|X_n| \leqslant Y$,Y 可积,那么 X 可积且

$$EX_n \to EX(n \to \infty)$$

[这仅仅是第一章 Lebesgue 控制收敛的不同表达式,可参见习题 1.7,15]

习　题　2.3

1. 设 ξ_1, ξ_2 都是 $(\Omega, \mathscr{F}, P)$ 上的随机变量,随机向量 (ξ_1, ξ_2) 的分

布函数定义为

$$F(x_1,x_2)=P(\xi_1\leqslant x_1,\xi_2\leqslant x_2)$$

证明下列两个函数：

$$G(x,y)=\begin{cases}1, & \text{当 } x+y\geqslant 0\\ 0, & \text{当 } x+y<0\end{cases}$$

$G(x,y)=[x+y]$,[·]表示整数部分关于每个变量均右连续，但不是分布函数。

2. 定义

$$\rho=\frac{\mathrm{Cov}(X_1,X_2)}{\sqrt{\mathrm{Var}X_1\cdot\mathrm{Var}X_2}}$$

令$\sigma_i=+\sqrt{\mathrm{Var}X_i}$, $EX_i=m_i$, $i=1,2$. 写出用 $(m_1, m_2, \sigma_1, \sigma_2, \rho)$ 参数表示的二维正态概率密度.

2.4 条件数学期望与一致可积性

提要 条件数学期望与一致可积性是学习鞅论的直接基础。因此，学好本节内容是学好以下各节的前提条件。本节给出了条件数学期望的严格定义及存在性证明，并讨论了条件数学期望的运算性质以及 Jensen 不等式。本节还给出了一致可积性的充分必要条件和几乎处处收敛序列 L_1 收敛的充分必要条件。

本节将给出在 σ 代数条件下的条件数学期望，这是一个相当抽象的概念。为了引出这一抽象概念，我们先从复习初等概率中的条件概率着手，然后讨论离散随机变量下的条件数学期望，最后定义一般 σ 代数下的条件数学期望并证明其存在性。这是一个由特殊到一般的抽象过程，读者应该反复阅读这些内容，仔细对照领会，才能真正掌握其实质。

1. 事件 B 条件下的条件数学期望

设掷一颗骰子出现四点为事件 A，由于骰子的均匀性和对称性，事件 A 出现的概率为 1/6，即

$$P(A)=\frac{1}{6}$$

设掷一颗骰子出现偶数点为事件 B，在 B 出现条件下 A 出现的概率称为条件概率，记作

$$P(A|B)=\frac{1}{6}\left(\frac{1}{2}\right)^{-1}=\frac{1}{3}$$

在一般概率空间 $(\Omega, \mathscr{F}, P)$ 上，$A,B\in\mathscr{F}$，在事件 B 出现的条件下 A 出现的条件概率为

$$P(A|B)=\frac{P(A\cap B)}{P(B)},\quad \forall P(B)\neq 0 \qquad (2.4.1)$$

当然，这里假定 $P(B)\neq 0$。

对于给定的 B，$P(A|B)$ 可作为 A 的集函数，它具有下列性质：

(1) $P(\Omega|B)=1$；

(2) $P(A|B)\geqslant 0, \forall A\in\mathscr{F}$；

(3) 若 $\{A_n\}$ 为两两不相交的事件序列，那么

$$P\left(\sum_{i=1}^{\infty} A_i\,\Big|\,B\right)=\sum_{i=1}^{\infty} P(A_i|B)$$

由此可见 $P(\cdot|B)$ 是可测空间 $(\Omega,\mathscr{F})$ 上的一个概率测度。

式(2.4.1)也可表示为

$$P(A|B)=\frac{\int_B \chi_A dP}{P(B)} \qquad (2.4.2)$$

我们把式(2.4.2)记作

$$P(A|B)=E[\chi_A|B] \tag{2.4.3}$$

设

$$\bigcup_{k=1}^{n} E_k=\Omega,\quad \varphi(\omega)=\sum_{k=1}^{n} a_k\chi_{E_k}$$

那么 $E[\varphi|B]$ 定义为

$$E[\varphi|B]\triangleq \frac{\sum\limits_{k=1}^{n} a_k\int_B \chi_{E_k}dP}{P(B)}=\frac{\int_B \varphi dP}{P(B)}$$

由于非负可测函数必为一简单函数列的极限(引理1.2.5),因此对于非负可测函数条件数学期望的定义为

$$E[X(\omega)|B]\triangleq\frac{\int_B XdP}{P(B)} \tag{2.4.4}$$

而任意 $\mathscr{F}$ 可测的函数 $X(\omega)$ 可写成 $X=X^+-X^-$, X^+ 和 X^- 均为非负可测函数,因此对于任意 $\mathscr{F}$ 可测函数 $X(\omega)$,在条件 B 下的条件数学期望可以作与(2.4.4)式同样形式的定义,但要求 $X(\omega)$ 可积。

2. 条件为离散型随机变量构成的 σ 代数

设 $(\Omega,\mathscr{F},P)$ 为概率空间,f 是 Ω 到 $N=\{1,2,\cdots\}$ 的可测映照,即 f 是离散型随机变量

$$\Omega\xrightarrow{f}N$$

令

$$A_n=\{\omega|f(\omega)=n\}\in\mathscr{F},\quad \forall=1,2,\cdots$$

$X(\omega)$ 是 $(\Omega,\mathscr{F},P)$ 上的可积随机变量。在 A_n 条件下,X 的条件数学期望为

$$\alpha_n = \frac{\int_{A_n} X dP}{P(A_n)}, \quad \forall P(A_n) \neq 0$$

这里$\int_{A_n} XdP$是可测函数X在A_n上关于测度P的积分，$P(A_n)$是A_n出现的概率。在已给离散型随机变量$f(\omega)$下，X的条件数学期望定义为

$$Y(\omega) = \sum_{k=1}^{\infty} \alpha_k \chi_{A_k} \tag{2.4.5}$$

这里χ_{A_k}是可测集A_k的示性函数，且

$$\alpha_k = \frac{\int_{A_k} X dP}{P(A_k)}, \quad \forall P(A_k) \neq 0$$

对于$P(A_k)$等于零的集，$Y(\omega)$可取任意值。因为$A_k \in \mathscr{F}$，$k = 1, 2, \cdots$，$Y(\omega)$是$\mathscr{F}$可测函数，因此其条件数学期望不是一个数值而是一个随机变量。

3. 一般情形

设$(\Omega, \mathscr{F}, P)$是概率空间，令$E = \Omega$，$\mathscr{E}$是$\mathscr{F}$的完备子σ代数，$X(\omega)$是概率空间$(\Omega, \mathscr{F}, P)$上的随机变量，且

$$\int_{\Omega} |X| dP < \infty$$

那么，存在定义在E上且满足如下条件的实可测函数($\mathscr{F}$可测函数)$Y(\omega)$：

(1) $Y(\omega)$关于$\mathscr{E}$可测[注意，$X(\omega)$不一定关于$\mathscr{E}$可测]；

(2) 对于任意的$A \in \mathscr{E}$

$$\int_A X dP = \int_A Y dP \tag{2.4.6}$$

$Y(\omega)$ 在几乎处处意义下唯一确定。

事实上，关于 $\mathscr{B}$ 的可测函数 $Y(\omega)$ 是存在的。由于 X 可积，$\int_A XdP$ 是 $(\Omega,\mathscr{B})$ 上的有限测度，且关于 P 绝对连续。因此 $(\Omega,\mathscr{B},\mu)$ 组成一测度空间，测度 μ 定义为

$$\mu(A)=\int_A XdP,\quad \forall A\in\mathscr{B}$$

$\mu\ll P$。根据 Radon-Nikodym 定理，满足式 (2.4.6) 的 $Y(\omega)$ 必定存在，且使

$$\mu(A)=\int_A Y(\omega)dP$$

这里

$$Y(\omega)=\frac{d\mu}{dP},\ \text{a.s.}$$

$Y(\omega)$几乎处处唯一确定。我们称这样的 $Y(\omega)$ 是在 $\mathscr{B}$ 代数下的条件数学期望，它是 μ 关于 P 的 Radon-Nikodym 导数，是一个随机变量。为形象起见，我们把 $Y(\omega)$ 记作

$$Y(\omega)=E(X|\mathscr{B})$$

对于 2. 中提到的离散型随机变量 f，$\mathscr{B}$ 是由离散事件组成的 σ 代数，对于 $A_n\in\mathscr{B}$，由式 (2.4.5) 可知

$$\int_{A_n}Y(\omega)dP=\int_{A_n}\sum_i a_i\chi_{A_i}dP$$

$$=a_nP(A_n)=\int_{A_n}XdP$$

这一等式是式 (2.4.6) 的特例。因此以离散事件 σ 代数为条件的条件数学期望是一般的 σ 事件代数为条件的条件数学期望的特殊情形。一般情形下的条件数学期望 $Y(\omega)$，是离散型随机变量条件下条件数学期望的推广。于是，我们可以作如下定义：

定义 2.4.1 假设 $X(\omega)$ 是定义在概率空间 $(\Omega, \mathscr{F}, P)$ 上的实值可积随机变量，$\mathscr{E}$ 是 $\mathscr{F}$ 的子 σ 代数，在给定 $\mathscr{E}$ 下，$X(\omega)$ 的条件数学期望是 $\mathscr{E}$ 可测的可积随机变量 $Y(\omega)$，它使得

$$\int_A X(\omega)dP = \int_A Y(\omega)dP, \quad \forall A \in \mathscr{E} \tag{2.4.7}$$

成立。

$Y(\omega)$ 记作 $E(X|\mathscr{E})$，称为 σ 代数 $\mathscr{E}$ 条件下，X 的条件数学期望。

式 (2.4.7) 可写成

$$\int_A X(\omega)dP = \int_A E(X|\mathscr{E})dP, \quad \forall A \in \mathscr{E} \tag{2.4.8}$$

若 $\mathscr{E}$ 是由随机变量族 f_i，$i \in I$ 产生的 σ 代数，即 $\mathscr{E} = \sigma\{f_i, i \in I\}$，则 $Y(\omega)$ 也称作在已给 f_i 下的条件数学期望，并记作

$$E(X|\mathscr{E}) = E(X|f_i, i \in I)$$

若 $X(\omega) = \chi_B(\omega)$，$B \in \mathscr{E}$，则 $Y(\omega)$ 称为已给 $\mathscr{E}$ 下 B 的条件概率，记作 $P(B|\mathscr{E})$，由式 (2.4.8) 可得

$$P(B \cap A) = \int_A P(B|\mathscr{E})dP$$

$$P(B) = \int_\Omega P(B|\mathscr{E})dP \tag{2.4.9}$$

注意：这里的 $Y(\omega)$ 不是一个数，而是一个随机变量。$P(B|\mathscr{E})$ 也是一个随机变量。

下面叙述条件数学期望的性质：

设 $(\Omega, \mathscr{F}, P)$ 是概率空间，下面的随机变量都定义在它上面。$\mathscr{E}$ 是 $\mathscr{F}$ 的任一完备子 σ 代数。

引理 2.4.1 假设 X 和 Y 是可积随机变量，$\alpha, \beta, \gamma \in R$ 都是常数，那么

$$E[\alpha X+\beta Y+\gamma|\mathscr{G}]=\alpha E[X|\mathscr{G}]+\beta E[Y|\mathscr{G}]+\gamma, \text{ a.s.}$$

证明 根据定义 2.4.1，下列各式对于一切 $A\in\mathscr{G}$ 均成立，即

$$\int_A E(\alpha X+\beta Y+\gamma|\mathscr{G})dP=\int_A(\alpha X+\beta Y+\gamma)dP, \quad \forall A\in\mathscr{G}$$

$$=\alpha\int_A XdP+\beta\int_A YdP+\gamma\int_A dP$$

$$=\alpha\int_A E(X|\mathscr{G})dP+\beta\int_A E(Y|\mathscr{G})dP+\gamma\int_A dP$$

$$=\int_A[\alpha E(X|\mathscr{G})+\beta E(Y|\mathscr{G})+\gamma]dP, \quad \forall A\in\mathscr{G}$$

根据 1.4 节，可以证明：若

$$\int_A fdP=\int_A gdP, \quad \forall A\in\mathscr{G} \text{ 成立}$$

那么

$$f=g, \text{ a.s.}（见习题 2.4,1）$$

因此

$$E(\alpha X+\beta Y+\gamma|\mathscr{G})=\alpha E(X|\mathscr{G})+\beta E(Y|\mathscr{G})+\gamma, \text{a.s.} \quad ▮$$

引理 2.4.2 假设 X 和 Y 是可积随机变量，$X\leqslant Y$, a.s. 那么

$$E(X|\mathscr{G})\leqslant E(Y|\mathscr{G}), \text{ a.s.}$$

证明 由于 $X\leqslant Y$, a.s.，有

$$Y-X\geqslant 0, \text{ a.s.}$$

可知

$$\int_A(Y-X)dP\geqslant 0, \quad \forall A \text{ 成立}$$

因此有

$$\int_A E[(Y-X)|\mathscr{G}]dP\geqslant 0, \quad \forall A\in\mathscr{G}$$

由引理 2.4.1 可知

$$\int_A [E(Y|\mathscr{F}) - E(X|\mathscr{F})]dP \geqslant 0, \quad \forall A \in \mathscr{F}$$

根据推论 1.4.3 可得

$$E(X|\mathscr{F}) \leqslant E(Y|\mathscr{F}), \text{ a.s.} \quad \blacksquare$$

引理 2.4.3 (1) 假定 $\{X_n\}$, $n \in \boldsymbol{N}$ 是非负单调增随机变量序列, $X_n \uparrow X$, X 可积, 那么

$$E(X|\mathscr{F}) = \lim_n E(X_n|\mathscr{F}), \text{ a.s.} \tag{2.4.10}$$

(2) 设 $Z \leqslant X_n \uparrow X$ a.e., 且 $E|Z| < \infty$, $E|X| < \infty$, 那么

$$\lim_n E(X_n|\mathscr{F}) = E(X|\mathscr{F})$$

证明 (1) 因为 $0 \leqslant X_n \uparrow X$, 根据引理 2.4.2, 知 $0 \leqslant E(X_1|\mathscr{F}) \leqslant E(X_2|\mathscr{F}) \leqslant \cdots$, a.e. 所以对几乎所有的 ω 极限 $\lim\limits_{n\to\infty} E(X_n|\mathscr{F})$ 存在, 对极限不存在的 ω 我们定义为零, 经过这样补充定义后的极限是可测的. 运用单调收敛定理:

$$\int_A \lim E(X_n|\mathscr{F})dP = \lim \int_A E(X_n|\mathscr{F})dP$$

及

$$\lim \int_A X_n dP = \int_A \lim X_n dP = \int_A X dP = \int_A E(X|\mathscr{F})dP$$

再根据定义 2.4.1

$$\lim \int_A X_n dP = \lim \int_A E(X_n|\mathscr{F})dP \tag{2.4.11}$$

即得

$$\int_A \lim E(X_n|\mathscr{F})dP = \int_A E(X|\mathscr{F})dP$$

因此

$$\lim_n E(X_n|\mathscr{F}) = E(X|\mathscr{F}), \text{ a.s.}$$

(2) 显然, $0 \leqslant X_n - Z \uparrow X - Z$, a.s. 且 $E|X - Z| < \infty$,

由（1）知

$$\lim_{n\to\infty} E(X_n - Z|\mathscr{E}) = E(X - Z|\mathscr{E}), \text{ a.s.}$$

由此得

$$\lim_n E(X_n|\mathscr{E}) = E(X|\mathscr{E}), \text{ a.s.}$$ ∎

引理 2.4.4 假设 Y 是可积随机变量，X 是 $\mathscr{E}$ 可测函数，XY 可积，于是

$$E(XY|\mathscr{E}) = XE(Y|\mathscr{E}), \text{ a.s.} \tag{2.4.12}$$

证明 显然 $XE(Y|\mathscr{E})$ 是 $\mathscr{E}$ 可测函数。根据条件数学期望的定义

$$\int_A E(XY|\mathscr{E})dP = \int_A XYdP, \quad \forall A \in \mathscr{E}$$

所以只要证明了

$$\int_A XE(Y|\mathscr{E})dP = \int_A XYdP, \quad \forall A \in \mathscr{E}$$

就证明了式（2.4.12）。

先证 $X(\omega)$ 是 $\mathscr{E}$ 可测集 B 的示性函数的情形：

$$X(\omega) = \chi_B(\omega), \quad B \in \mathscr{E}$$

对于一切 $A \in \mathscr{E}$ 有

$$\int_A \chi_B E(Y|\mathscr{E})dP = \int_{A\cap B} E(Y|\mathscr{E})dP$$

$$= \int_{A\cap B} YdP \quad (\text{由定义 2.4.1 知}, A\cap B \in \mathscr{E})$$

$$= \int_A \chi_B YdP = \int_A XYdP$$

所以式（2.4.12）在 $X(\omega)$ 是 $\mathscr{E}$ 可测集的示性函数时成立。

现在证明简单函数的情形：

设

$$X(\omega) = \sum_{i=1}^{n} a_i \chi_{B_i}$$

其中 α_i 是实数，$B_i \in \mathscr{G}$ 且两两不相交。

$$\sum_{i=1}^{n} B_i = \Omega$$

对于任何 $A \in \mathscr{G}$ 有

$$\int_A XE(Y|\mathscr{G})dP = \sum_{i=1}^{n} \alpha_i \int_A \chi_{B_i} E(Y|\mathscr{G})dP$$

$$= \sum_{1}^{n} \alpha_i \int_{A\cap B_i} E(Y|\mathscr{G})dP$$

$$= \sum_{1}^{n} \alpha_i \int_{A\cap B_i} YdP$$

$$= \int_A \left(\sum_{1}^{n} \alpha_i \chi_{B_i}\right) YdP = \int_A XYdP$$

所以当 $X(\omega)$ 是简单函数时，式(2.4.12)成立。

设 $X_n^+ \geqslant 0$ 是一单调增简单函数列

$$\lim X_n^+ = X^+$$

故 $X_n^+Y^+$ 也是单调增序列，且

$$\lim X_n^+Y^+ = X^+Y^+$$

由引理 2.4.3 可推得

$$E(X^+Y^+|\mathscr{G}) = \lim E(X_n^+Y^+|\mathscr{G})$$

$$= \lim X_n^+E(Y^+|\mathscr{G}) = X^+E(Y^+|\mathscr{G})$$

同理可得

$$E(X^+Y^-|\mathscr{G}) = X^+E(Y^-|\mathscr{G})$$

因此

$$E(X^+Y|\mathscr{G}) = X^+E(Y|\mathscr{G})$$

同理可证 $E(X^-Y|\mathscr{G}) = X^-E(Y|\mathscr{G})$，故式(2.4.12)成立。▌

引理 2.4.5 若 Y 是可积随机变量，$\mathscr{G}_1, \mathscr{G}_2$ 是 $\mathscr{F}$ 的两个

子σ代数，如果 $\mathscr{E}_1\subset\mathscr{E}_2\subset\mathscr{F}$，那么

$$E\{E(Y|\mathscr{E}_2)|\mathscr{E}_1\}=E(Y|\mathscr{E}_1) \qquad (2.4.13)$$

$$E\{E(Y|\mathscr{E}_1)|\mathscr{E}_2\}=E(Y|\mathscr{E}_1) \qquad (2.4.14)$$

证明 因为 $E(Y|\mathscr{E}_1)$ 关于 $\mathscr{E}_1$ 可测，$\mathscr{E}_1\subset\mathscr{E}_2$，所以 $E(Y|\mathscr{E}_1)$ 关于 $\mathscr{E}_2$ 也可测．由引理 2.4.4，有

$$E\{E(Y|\mathscr{E}_1)|\mathscr{E}_2\}=E(Y|\mathscr{E}_1)$$

即式 (2.4.14) 成立．

证明式 (2.4.13)：因为式 (2.4.13) 的两边关于 $\mathscr{E}_1$ 均可测，故对 $A\in\mathscr{E}_1$ 有

$$\int_A E\{E(Y|\mathscr{E}_2)|\mathscr{E}_1\}dP=\int_A E(Y|\mathscr{E}_2)dP$$

$$=\int_A YdP=\int_A E(Y|\mathscr{E}_1)dP$$

所以

$$E\{E(Y|\mathscr{E}_2)|\mathscr{E}_1\}=E(Y|\mathscr{E}_1), \text{ a.s.} \qquad \blacksquare$$

引理 2.4.6 假如 $\mathscr{D}$ 是一个平凡σ代数，即 $\mathscr{D}=(\Omega,\varnothing)$，$E(X|\mathscr{D})=E(X)$，那么

$$E(E(X|\mathscr{E})|\mathscr{D})=E[E(X|\mathscr{E})]=E(X)$$

证明 因为 $\mathscr{D}\subset\mathscr{E}$，由引理 2.4.5 有

$$E(E(X|\mathscr{E})|\mathscr{D})=E(X|\mathscr{D})=E(X) \qquad \blacksquare$$

引理 2.4.7 假设 X 是可积随机变量，$\mathscr{E}$ 是 $\mathscr{F}$ 的子σ代数，$X=E(X|\mathscr{E})$，a.s. 的充分必要条件是 X 为 $\mathscr{E}$ 可测函数．

证明 若 $E(X|\mathscr{E})=X$，由 $E(X|\mathscr{E})$ 的定义可知，$E(X|\mathscr{E})$是 $\mathscr{E}$ 可测函数，X 为 $\mathscr{E}$ 可测．反之，若 X 是 $\mathscr{E}$ 可测，由于

$$\int_A XdP=\int_A E(X|\mathscr{E})dP, \ \forall A\in\mathscr{E}$$

所以

$$X = E(X \mid \mathscr{G}), \text{ a.s.}$$ ∎

为了方便讨论条件数学期望的不等式，这里先给出凸函数的定义。

定义 2.4.2 φ 是 R 上的连续实值函数，若对于任何 $x_1, x_2 \in R$，有

$$\varphi\left(\frac{x_1 + x_2}{2}\right) \leqslant \frac{\varphi(x_1) + \varphi(x_2)}{2}$$

则称 φ 是凸函数。

φ 的图像如图 2.4.1 所示。

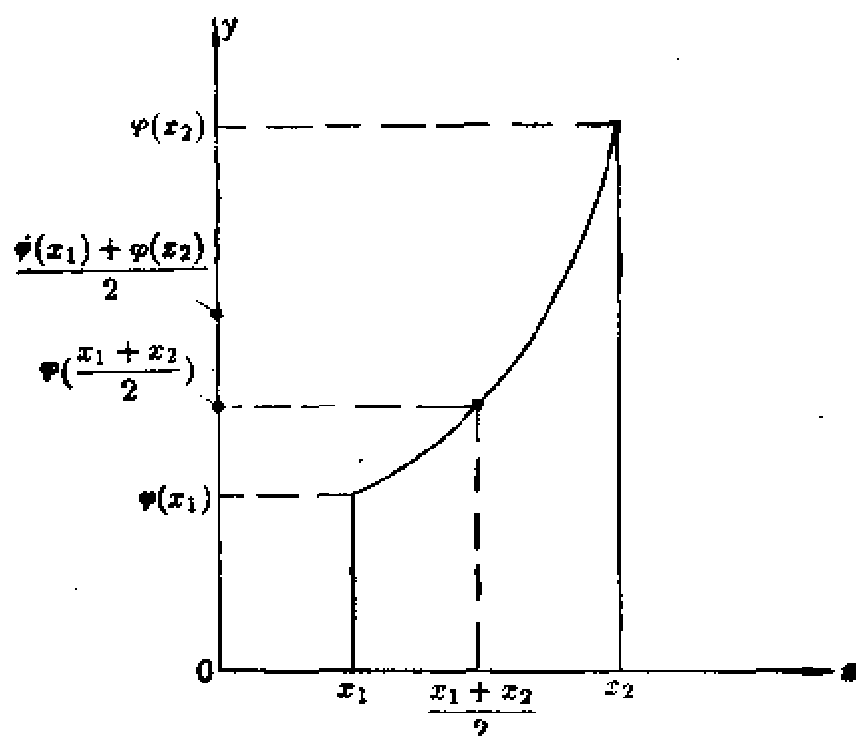


图 2.4.1 凸函数示意图

一般地，$\phi(\cdot)$ 是 $R \to R$ 的一个映照，若对任何 $\lambda_1, \cdots, \lambda_n$，$\sum_{i}^{n} \lambda_i = 1$，且

$$\phi\left(\sum_{i=1}^{n} \lambda_i x_i\right) \leqslant \sum_{i=1}^{n} \lambda_i \phi(x_i)$$

成立，则称 $\phi(\cdot)$ 为凸映照。

引理 2.4.8（Jensen 不等式） 假设 ϕ 是 R 到 R 的连续凸映照，X 是可积随机变量且 $\phi(X)$ 可积，那么

$$\phi(E(X|\mathscr{G}))\leqslant E(\phi(X)|\mathscr{G}),\ \text{a.s.}\tag{2.4.15}$$

证明 对任何 x,y 有

$$\phi(x)-\phi(y)\geqslant\phi'(y)(x-y)$$

其中 ϕ' 是 ϕ 的右导数。这是凸函数的一个特性。令 $x=X$, $y=E(X|\mathscr{G})$，有

$$\phi(X)-\phi(E(X|\mathscr{G}))\geqslant\phi'(E(X|\mathscr{G}))[X-E(X|\mathscr{G})]$$

两边取条件数学期望

$$\begin{aligned}&E\{[\phi(X)-\phi(E(X|\mathscr{G}))]|\mathscr{G}\}\\ \geqslant\ &\phi'(E(X|\mathscr{G}))E\{[X-E(X|\mathscr{G})]|\mathscr{G}\}=0\end{aligned}$$

得

$$E(\phi(X)|\mathscr{G})\geqslant\phi(E(X|\mathscr{G}))\ \text{a.s.}$$

（这里要求 ϕ' 是 $\mathscr{G}$ 可测的，见习题 2.4,9） ▮

取一个具体的凸函数：

$$\phi(x)=|x|^p,\ 1\leqslant p<\infty$$

利用 Jensen 不等式可得到

$$\|E(X|\mathscr{G})\|_p\leqslant\|X\|_p,\ \ \forall X\in L^p(\Omega,\mathscr{F},P)$$

这里

$$\|E(X|\mathscr{G})\|_p=\left(\int_\Omega|E(X|\mathscr{G})|^p dP\right)^{1/p}$$

而

$$\|X\|_p=\left(\int_\Omega|X|^p dP\right)^{1/p}$$ ▮

一致可积性概念与条件数学期望概念没有直接关系。但它是鞅论中十分重要的概念。我们放在本节讨论，以便从下一节开始讨论鞅论。

$(\Omega,\mathscr{F},P)$ 是概率空间，$\|X\|_1$ 定义为 $|X|$ 的积分，即

$$\|X\|_1 = E|X|$$

$$L^1(\Omega, \mathscr{F}, P) \triangleq \{X : \|X\|_1 < \infty\}$$

$L^1(\Omega, \mathscr{F}, P)$ 是一个函数空间，称为 L^1 空间。

定义 2.4.3 假设 K 是 L^1 的一个子集，即

$$K \subset L^1(\Omega, \mathscr{F}, P)$$

若 $c \to \infty$ 时

$$\int_{|X| \geq c} |X| dP \to 0, \quad \forall X \in K \text{ 一致成立}$$

即对于任给 $\varepsilon > 0$，存在一个 c_0(c_0 与 X 无关，$\forall X \in K$)；若 $c > c_0$ 时

$$\int_{|X| \geq c} |X| dP < \varepsilon, \quad \forall X \in K$$

则称 K 是 $L^1(\Omega, \mathscr{F}, P)$ 中的一个一致可积子集。

若 $K = \{X_n, n \geq 1\}$，即 K 是 $L^1(\Omega, \mathscr{F}, P)$ 中一个随机变量序列，则称 $\{X_n\}$ 一致可积，或 $\{X_n\}$ 具有一致可积性。

若 X 是随机变量，$c > 0$，定义

$$X^c(\omega) = \begin{cases} X(\omega), & \text{当 } |X(\omega)| < c \\ 0, & \text{当 } |X(\omega)| \geq c \end{cases}$$

令

$$X_c(\omega) = X(\omega) - X^c(\omega)$$

K 一致可积的充分必要条件，是对任何 $\varepsilon > 0$，存在 $c > 0$ 使得

$$\|X_c\|_1 < \varepsilon, \forall X \in K \subset L^1(\Omega, \mathscr{F}, P)$$

成立。事实上，由于

$$|X(\omega)| = |X_c(\omega)| + |X^c(\omega)|$$

$$\int_{|X| \geq c} |X(\omega)| dP = \int_{|X| \geq c} (|X_c(\omega)| + |X^c(\omega)|) dP$$

因此

$$\int_{|X| \geq c} |X| dP = \int_{|X| \geq c} |X_c| dP = \|X_c\|_1 < \varepsilon$$

$$\forall X \in K \text{ 一致成立}$$

定理2.4.1 假设 $K \subset L^1(\Omega, \mathscr{F}, P)$，$K$ 是一致可积的充分必要条件，是下面两个条件同时成立：

(1) 存在一个有限数 a，即 $a < \infty$，使得

$$E|X| < a, \quad \forall X \in K \text{ 成立}$$

即 $E|X|$ 在 K 上有界；

(2) 对任何 $\varepsilon > 0$，存在一个 $\delta > 0$，当

$$P(A) \leqslant \delta, \quad \forall A \in \mathscr{F}$$

时，有

$$\int_A |X| dP < \varepsilon, \quad \forall X \in K \text{ 成立}$$

证明 必要性：

对于任何一个可积随机变量 X 和任何一个 $A \in \mathscr{F}$，均有

$$\int_A |X| dP \leqslant cP(A) + E|X_c|$$

因为 K 一致可积，对任给 $\varepsilon > 0$ 必存在一个 $c > 0$ 使得

$$E|X_c| < \varepsilon/2, \quad \forall X \in K \text{ 成立}$$

于是

$$E|X| \leqslant c + \varepsilon/2, \quad \forall X \in K$$

因此条件 (1) 满足。

对于同样的 c，如果 $P(A) \leqslant \delta = \varepsilon/2c$，那么

$$\int_A |X| dP \leqslant cP(A) + E|X_c|$$

$$< c \cdot \frac{\varepsilon}{2c} + \frac{\varepsilon}{2} = \varepsilon, \quad \forall X \in K$$

因此条件 (2) 满足。

充分性：

假设条件 (1) 和 (2) 满足，对任给 $\varepsilon > 0$，存在 $\delta > 0$，取

$$c = \delta^{-1}\sup_{X\in K} E|X|$$

因为 $E|X|$ 在 K 上有界，所以上确界必存在。

设 $A_X = \{|X| \geqslant c\}$，则

$$P(A_X) = P\{|X| \geqslant c\} \leqslant c^{-1}E|X| \leqslant \delta$$

于是

$$\int_{\{|X|\geqslant c\}} |X(\omega)|dP = \int_{A_X} |X(\omega)|dP < \varepsilon, \quad \forall\ X \in K \text{ 成立}$$

故 K 一致可积。 ∎

推论 2.4.1 $K \subset L^1(\Omega, \mathscr{F}, P)$。 假设 $\phi(t)$ 是定义在 $[0,\infty)$ 上的正增函数且

$$\lim_{t\to\infty} \frac{\phi(t)}{t} = +\infty$$

$$\sup_{X\in K} E\phi(|X|) < \infty$$

那么 K 一致可积。

证明 设 $S = \sup\limits_{X\in K} E\phi(|X|)$，考虑 $\varepsilon > 0$，令

$$a = \varepsilon^{-1}S$$

选择足够大的 c 使得

$$\frac{\phi(t)}{t} \geqslant a, \text{ 当 } t \geqslant c$$

在集合 $\{|X| \geqslant c\}$ 上

$$|X| \leqslant a^{-1}\phi(|X|)$$

因此

$$\int_{\{|X|\geqslant c\}} |X(\omega)|dP \leqslant a^{-1}\int_{\{|X|\geqslant c\}} \phi(|X|)dP$$
$$\leqslant a^{-1}E\phi(|X|) \leqslant \varepsilon, \forall X \in K$$

所以 K 是一致可积的。 ∎

下面的定理将说明，$\{X_n\}$ 一致可积，那么，几乎处处收敛必 L^1 收敛。

定理2.4.2 假设 $\{X_n\}$ 是可积随机变量序列，$\lim X_n = X$，a.s. X 可积，那么 $\{X_n\}$ 必 L^1 收敛，即

$$\lim_{n\to\infty}\int_{\Omega}|X_n - X|\,dP = 0$$

的充分必要条件是 $\{X_n\}$ 一致可积。

证明 必要性：

由于对任何 $A\in\mathscr{F}$ 有

$$\int_A |X_n|\,dP \leqslant \int_A |X|\,dP + \|X_n - X\|_1 \qquad (2.4.16)$$

取 $A=\Omega$，对于任何 $\varepsilon>0$，存在 N，当 $n>N$ 时

$$\|X_n - X\|_1 < \varepsilon/2$$

所以

$$E|X_n| \leqslant \int |X|\,dP + \max\Big(\|X_1 - X\|_1,\cdots,$$

$$\|X_N - X\|_1,\ \frac{\varepsilon}{2}\Big) = a$$

因此 $E|X_n|\leqslant a$ 一致有界，满足定理 2.4.3 中条件 (1)。

考虑一个个数有限的随机变量族：

$$\mathscr{S} = \{X_1 - X, X_2 - X, \cdots, X_N - X, X\}$$

由于它们是有限个可积随机变量，因此总存在 $\delta>0$ 使得 $A\in\mathscr{F}$，$P(A)\leqslant\delta$ 时，有

$$\int_A |Y|\,dP \leqslant \varepsilon/2,\ \forall Y\in\mathscr{S}$$

再根据式 (2.4.16)，对 $A\in\mathscr{F}$，$P(A)\leqslant\delta$，有

$$\int_A |X_n|\,dP < \varepsilon,\ \forall X_n \text{ 成立}$$

即满足定理 2.4.1 中条件 (2)。$\{X_n\}$ 满足定理 2.4.1 中条件 (1)，(2)，所以 $\{X_n\}$ 一致可积。

充分性：

假设 $\{X_n\}$ 一致可积，那么存在常数 a，

$$E|X_n| < a, \forall X_n \text{ 成立}$$

根据 $\lim X_n = X$, a.s. 知 $\lim |X_n| = |X|$, a.s.。

由 Lebesgue 控制收敛定理

$$\lim_n E(|X_n|) = E|X| \qquad (2.4.17)$$

由于

$$\|X_n - X\|_1 \leqslant \|X_n^c - X^c\|_1 + \|X_{n,c}\|_1 + \|X_c\|_1$$

而且 $\{X_n\}$ 一致可积，因此对任何 $\varepsilon > 0$，存在 $c > 0$ 使得

$$\|X_{n,c}\|_1 < \varepsilon/3$$

再由 X 可积得

$$\|X_c\|_1 < \varepsilon/3$$

又因为

$$\lim |X_n^c - X^c| = 0, \text{ a.s.}$$

且

$$|X_n^c - X^c| \leqslant 2c$$

根据 Lebesgue 控制收敛定理

$$\lim_n \|X_n^c - X^c\|_1 = 0$$

因此

$$\|X_n^c - X^c\|_1 < \varepsilon/3$$

当 $n \to \infty$ 时，$\|X_n - X\|_1 \to 0$。

习　题　2.4

1. 若 f, g 是 $(\Omega, \mathscr{A}, P)$ 上的可测实函数，

$$\int_A f dP = \int_A g dP, \quad \forall A \in \mathscr{A}$$

证明 $f = g$ a.e.。

2. 第一章单调收敛定理 1.4.1 的推广形式：若 $\{f_n\}$ 是可积随机变量序列，$g \leqslant f_n \leqslant f_{n+1}$ $(n = 1, 2, \cdots,)$ $\lim f_n = f$，g 和 f 是可积随机变量，证明

$$\lim_n \int f_n dP = \int \lim f_n dP$$

3.设ξ与η是独立同分布随机变量，$E\xi<\infty$，证明

$$E(\xi|\xi+\eta)=E(\eta|\xi+\eta)=\frac{\xi+\eta}{2}, \text{ a.s.}$$

4.设$\xi_1, \xi_2, \cdots$是独立同分布随机变量序列且$E|\xi_i|<\infty$；$i=1, 2, \cdots$，证明

$$E(\xi_1|s_n, s_{n+1}, \cdots)=\frac{s_n}{n}, \text{ a.s.}$$

其中$s_n=\xi_1+\cdots+\xi_n$。

5.设ξ是具有分布函数$F_\xi(x)$的随机变量，证明

$$E(\xi|a<\xi\leqslant b)=\frac{\int_a^b x dF_\xi(x)}{F_\xi(b)-F_\xi(a)}$$

假定$F_\xi(b)-F_\xi(a)>0$。

6.设(ξ, η)是一对随机变量，ξ是能观察到的，η则不能，若$E\eta^2<\infty$，证明

$$\varphi^*(x)=E(\eta|\xi=x)$$

定义的$\varphi^*(\xi)$是η的最优估计，即

$$E[\eta-\varphi^*(\xi)]^2=\min_\varphi E[\eta-\varphi(\xi)]^2$$

7.证明概率论中的Fatou引理：设$\{\xi_n\}_{n\geqslant 1}$是一列非负一致可积的随机变量，$E\limsup\xi_n$可积，那么

$$\limsup_n E\xi_n\leqslant E\limsup\xi_n$$

8.设$\xi_n\leqslant\eta_n$，$n\geqslant 1$，其中$\{\xi_n^+\}$一致可积，$\lim_n\eta_n=\eta$，a.s. 证明

$$\limsup E\xi_n\leqslant E\limsup\xi_n$$

9.设φ是连续凸函数，X关于$\mathscr{F}$可测，证明其右导数$\varphi'(X)$关于$\mathscr{F}$可测。

2.5 适应、停时和鞅

提要 本节引入随机序列与σ代数增序列相适应的概念，以及停时、初遇时间的概念和有关定理。

还给出了鞅和鞅的等价定义。

概率空间是概率论中的一个基本的概念。然而对于与时间有关的随机现象，光有概率空间的概念还不够，还必须引入与时间有关的事件 σ 代数. 这对于今后的讨论是非常重要的.

以后，我们用 $(\Omega, \mathscr{F}, P)$ 表示完备概率空间，$(\mathscr{F}_n, n\in \boldsymbol{N})$ 表示 $\mathscr{F}$ 的完备子 σ 代数增序列。完备子 σ 代数增序列包括下列三层意义：

(1) $\mathscr{F}_n$ 是 σ 代数且 $\mathscr{F}_n \subset \mathscr{F}$, $n\geqslant 1$.

(2) $\mathscr{F}_n$ 包含了 $\mathscr{F}$ 中一切测度为零的 Ω 的子集，$n\geqslant 1$，即 $\mathscr{F}_n$ 是完备子 σ 代数.

(3) $\mathscr{F}_n \subset \mathscr{F}_{n+1}$, $n\geqslant 1$，即单调增序列. 记

$$\mathscr{F}_\infty = \bigvee_{n\in \boldsymbol{N}} \mathscr{F}_n$$

在滤波、控制等实际应用中

$$\mathscr{F}_n = \sigma(Y_1, \cdots, Y_n), \quad n\in \boldsymbol{N}$$

是一个自然的 σ 代数增序列，这里 $\sigma(Y_1, \cdots, Y_n)$ 表示 $(Y_1, \cdots, Y_n)$ 的自然 σ 代数或由 $(Y_1, \cdots, Y_n)$ 生成的 σ 代数，见 2.3 节中式 (2.3.2).

引进了 $(\Omega, \mathscr{F}, P)$ 上，$\mathscr{F}$ 的("完备"今后不再说明)子 σ 代数增序列 $\mathscr{F}_n$ 之后，我们把

$$(\Omega, \mathscr{F}, \mathscr{F}_n, P) \tag{2.5.1}$$

四元组称为滤波空间.

设 $(X_n, n\in \boldsymbol{N})$ 是定义在 $(\Omega, \mathscr{F}, \mathscr{F}_n, P)$ 上的随机变量序列，若 X_n 关于 $\mathscr{F}_n$ 可测，$n\geqslant 1$，则称 X_n 是适应的. 如果 $(X_n, n\in \boldsymbol{N})$ 是 $(\Omega, \mathscr{F}, \mathscr{F}_n, P)$ 上的适应序列，$f_n(\cdot)$ 是 R^n 上的 Borel 可测函数，那么

$$X_n' = f_n(X_1, \cdots, X_n), \quad n\in \boldsymbol{N}$$

也是适应序列. 若 $(X_n, n\in\bar{N})$序列 $\bar{N}$ 表示 n 可取值 $+\infty$，X_n 关于 $\mathscr{F}_n$ 可测，X_∞ 关于 $\mathscr{F}_\infty$ 可测，则称 $(X_n, n\in\bar{N})$ 是适应的.

停时是鞅论中一个重要的基本概念.

停时最初被解释为停止赌博[1]的时间. 为了说明停时是怎么一回事，下面简单叙述一个称为"加倍或输光"的赌博问题.

掷均匀硬币，可以用一列独立同分布的随机变量 $\{y_n\}_{n\geq 1}$ 来描述，

$$y_n = \begin{cases} 1, & \text{第 } n \text{ 次掷出正面} \\ -1, & \text{第 } n \text{ 次掷出反面} \end{cases} \tag{2.5.2}$$

掷出正面时赌徒赢，赌徒得到原赌本的一倍，掷出反面时赌徒输，其赌本全部输光. 如果赌徒从一元钱的赌本开始赌起，那么掷 n 次后赌徒的赌本为 x_n，

$$x_n = \prod_{i=1}^{n}(1+y_i) \tag{2.5.3}$$

由式(2.5.3)清楚地看出，只要某一时刻 k，$y_k=-1$ 则 $x_k=0$. 这就是说，赌徒应在一定的时间停止赌博，否则他迟早会输光. 停止赌博的时刻记作 τ，这时赌徒的赌本为 x_τ. 显然，τ 不是一个固定的时刻，它应是一个随机变量. 然而赌徒只能根据已经进行过的赌局来决定停止赌博，还是继续赌博，他不能预知还未进行的赌局的结果. 这就是说 $\tau\leqslant n$ 还是 $\tau>n$ 完全决定于 $y_1,\cdots,y_n$ 而与 $y_{n+1}, y_{n+2},\cdots$ 无关.

先考虑离散时间停时. 用数学的语言可以叙述如下：

定义 2.5.1 设 $(\Omega, \mathscr{F}, \mathscr{F}_n, P)$ 为一滤波空间

1）这里为说明鞅含义的方便和用语简洁，借用了"赌博"一词，并非研究或赞赏赌博.

$$\tau: \Omega \to \bar{N} \text{ 的映照} \tag{2.5.4a}$$

如果

$$\{\omega \mid \tau(\omega) \leqslant n\} \in \mathscr{F}_n, \forall n \in N \tag{2.5.4b}$$

则称 $\tau(\omega)$ 为停时。

下面的引理给出了停时的等价定义。

引理 2.5.1 设 $\nu: \Omega \to \bar{N}$ 的映照，ν 为停时的充分必要条件是：$\{\omega \mid \nu(\omega) = n\} \in \mathscr{F}_n, \forall n \in N$.

证明 因为

$$\{\nu \leqslant n\} = \bigcup_{m \leqslant n} \{\nu = m\}$$

又 $\mathscr{F}_1 \subset \mathscr{F}_2 \subset \cdots \subset \mathscr{F}_n \subset \cdots$，故引理中的条件 $\{\omega \mid \nu(\omega) = n\} \in \mathscr{F}_n$ 是充分的。

再看

$$\{\nu = n\} = \{\nu \leqslant n\} - \{\nu \leqslant n-1\}$$

由于 $\mathscr{F}_n$ 是递增序列，因此，这条件又是必要的。 ▮

我们可以把停时 ν 与子 σ 代数联系起来，$\mathscr{F}_\nu$ 是 Ω 的子集定义为

$$\mathscr{F}_\nu = \{B: B \in \mathscr{F}_\infty,\ B \cap \{\nu = n\} \in \mathscr{F}_n,\ n \in N\} \tag{2.5.5}$$

我们把 $\mathscr{F}_\nu$ 中的事件称为先于时刻 ν 的事件，$\mathscr{F}_\nu$ 是一个 σ 代数而且还是 $\mathscr{F}_\infty$ 的子 σ 代数。

停时的一个重要例子是初遇时间。

定义 2.5.2 设 $(X_n, n \in N)$ 是滤波空间 $(\Omega, \mathscr{F}, \mathscr{F}_n, P)$ 上的适应随机序列，它取值于 $(R, \mathscr{B})$，若 $F \in \mathscr{B}$

$$\nu_F(\omega) \triangleq \begin{cases} \inf\{n: X_n(\omega) \in F\}, & \omega \in \bigcup_n \{X_n \in F\} \\ +\infty, & \text{其他场合} \end{cases}$$

则 $\nu_F(\omega)$ 称为可测集 F 的初遇时间。 其物理意义为 $X_n(\omega)$

最初与集 F 相遇的时刻。

因为对任何 n，$n\in N$

$$\{\nu_F = n\} = \bigcap_{m<n}\{X_m\in F'\}\cap\{X_n\in F\}\in\mathscr{F}_n$$

这就是说

$$\nu_F:\Omega\to\bar{N}$$

由引理 2.5.1 知，ν_F 是停时。

显然，$\bar{N}$ 中的任何整数均可以看作停时。这就是说非负的整常数是停时。设 $P\in\bar{N}$，$\nu\equiv P$，P 为常数，那么 $\mathscr{F}_\nu=\mathscr{F}_P$。

如果与随机现象有关的时间 t 取值于 $\bar{R}_+$，即 t 可取非负实数和 $+\infty$。对此，我们可以作类似的讨论：

设 $\{\mathscr{F}_t\}_{t\in R_+}$ 是一族 σ 代数。如果 $\mathscr{F}_t$ 满足下列条件，则我们称它为一个满足通常条件的 σ 代数族：

(1) $\mathscr{F}_t$ 是递增的，即当 $s<t$ 时 $\mathscr{F}_s\subset\mathscr{F}_t$；

(2) $\mathscr{F}_t$ 是完备的，设 $A\subset\Omega$，A 是 $\mathscr{F}$ 的零测度集，则 $A\in\mathscr{F}_t$；

(3) $\mathscr{F}_t$ 右连续：

$$\mathscr{F}_t=\bigcap_{s>t}\mathscr{F}_s$$

在应用中，我们总假设 $\mathscr{F}_t$ 是时刻 t 以前发生的事件全体，所以 $\mathscr{F}_t$ 递增的假设是自然的，而条件 (2) 与 (3) 都是技术性的。

连续时间情形的停时定义如下：

定义 2.5.3 设 $(\Omega,\mathscr{F},\mathscr{F}_t,P)$ 为一滤波空间，τ 是 $\Omega\to\bar{R}_+$ 的映照。如果对任何 t

$$\{\omega\,|\,\tau(\omega)\leqslant t\}\in\mathscr{F}_t$$

则称 $\tau(\omega)$ 为停时。

下面的引理给出了停时的一些性质：

引理2.5.2 若 τ_1，τ_2 是停时，那么

$$\tau_1 \wedge \tau_2 \triangleq \min(\tau_1, \tau_2)$$

和
$$\tau_1 \vee \tau_2 \triangleq \max(\tau_1, \tau_2)$$
$$\tau + \tau_2$$

都是停时。

证明 对任何 $t \in R$

$$\{\omega | \tau_1 \wedge \tau_2 \leqslant t\} = \{\tau_1 \leqslant t\} \cup \{\tau_2 \leqslant t\} \in \mathscr{F}_t$$
$$\{\omega | \tau_1 \vee \tau_2 \leqslant t\} = \{\tau_1 \leqslant t\} \cap \{\tau_2 \leqslant t\} \in \mathscr{F}_t$$
$$\{\tau_1 + \tau_2 \leqslant t\} = \{\tau_1 = 0, \tau_2 = t\} \cup \{\tau_1 = t,$$
$$\tau_2 = 0\} \cup \Big(\bigcup_{\substack{a+b\leqslant t \\ a,b>0}} [\{\tau_1 < a\} \cap \{\tau_2 < b\}] \Big) \in \mathscr{F}_t$$

其中 a,b 是有理数。 ▌

定义 2.5.4 设 $(X_n, n \in \boldsymbol{N})$ 是 $(\Omega, \mathscr{F}, \mathscr{F}_n, P)$ 上的适应序列，τ 是停时，对任何 n

$$\{X_n \in B\} \cap \{\tau = n\} \in \mathscr{F}_n \tag{2.5.6}$$

称 X_τ 关于 $\mathscr{F}_\tau$ 可测，这里 B 是任一 Borel 可测集。若 $A \in \mathscr{F}$ 且 $A \cap \{\tau = n\} \in \mathscr{F}_n$ 则 $A \in \mathscr{F}_\tau$。

引理 2.5.3 设 S，T 为停时。

(1) 若 $A \in \mathscr{F}_S$，那么 $A \cap [S \leqslant T] \in \mathscr{F}_T$，$A \cap [S = T] \in \mathscr{F}_T$；

(2) 若 $S \leqslant T$，那么 $\mathscr{F}_S \subset \mathscr{F}_T$。

证明

(1) 根据式 (2.5.5) 的定义，当 $A \in \mathscr{F}_S$ 时有 $A \in \mathscr{F}_\infty$ 且 $A \cap \{S = n\} \in \mathscr{F}_n$，$n \in \boldsymbol{N}$。

因为 S,T 均为 $\mathscr{F}_\infty$ 可测函数，所以

$$A \cap [S \leqslant T] \in \mathscr{F}_\infty$$

且 $\forall\ n \in \boldsymbol{N}$ 有

$$\{A\cap[S\leqslant T]\}\cap[T=n]$$
$$=\bigcup_{i\leqslant n}(A\cap[S\leqslant i])\cap[T=n]\in\mathscr{F}_n$$

故

$$A\cap[S\leqslant T]\in\mathscr{F}_T\ [\text{定义式 (2.5.5)}]$$

同理可证

$$A\cap[S=T]\in\mathscr{F}_T$$

(2) 设 $A\in\mathscr{F}_S$，则

$$A\cap[T\leqslant n]=A\cap[S\leqslant T]\cap[T\leqslant n]\in\mathscr{F}_n$$

即 $A\in\mathscr{F}_T$，故 $\mathscr{F}_S\subset\mathscr{F}_T$. ∎

鞅的英文原名是 Martingale，它在英语中是一个多义词，它有三种不同的含义：1) 鞅具、马颔(一种马具)；2) 帆船斜撑杆的前支索；3)"费厄泼赖"或"公平"赌博. 据文献记载，一般认为 Martingale 是由法国数学家 Ville (1939) 首先提出的，后来 Doob 做了大量的研究工作，完成了今天称之为经典鞅论的那些研究工作. 关于鞅曾有过这样一个小故事：有一天 Doob 收到他以前的学生、著名数学家 P. Halmos 寄给他的一个包裹，打开一看是一条有黄铜环的漂亮皮带. Doob 不知道 Halmos 为什么要寄条有黄铜环的皮带给他. 经过询问，Halmos 告诉他，你收到了一个 martingale (鞅).后来这条皮带成了 Doob 的一个珍藏品.

因此 Martingale 确有鞅的意思，我们国内就采用了这一译法. 但在概率论中，其真实含义却是第三种意义，即 Martingale 表示"公平"赌博.

为了说明鞅的含义，下面介绍一个简单的例子. 设赌徒在第 n 次赌博时有赌资 X_n; $n=1,2,\cdots$ 下一次的输赢是 Δ_n. 如果下一次的输赢相互抵消，即赌徒下一次的赌资在已给现有赌资的条件下的条件数学期望恰好等于现在的赌资，

那么赌博是"公平的"。因为在 n 次赌博时赌资 X_n 是随机变量。设由 $(X_1,\cdots,X_n)$ 生成的 σ 代数，记作 $\mathscr{F}_n$，那么

$$E(X_{n+1}|\mathscr{F}_n)=\frac{1}{2}(X_n+\Delta_n)+\frac{1}{2}(X_n-\Delta)=X_n$$

于是称 X_n 为鞅。

由这个例子我们看到"公平的赌博"是 Martingale 的真实含义。但我们必须指出：鞅论发展的真正动力是生产技术和科学发展的需要，而不是赌博。现在提起的鞅，在人们脑子里出现的是一种特殊的随机过程，而不再是提出鞅论时的"公平"赌博。人们对于鞅感兴趣的原因：在于它是概率论、数理统计、系统辨识中常遇到的一种随机过程，它具有良好的性质。特别是鞅收敛定理、鞅型序列如渐近鞅（记作 Amart）等，在时间序列分析、系统辨识和自适应控制理论中显得十分重要。

下面我们给出鞅的数学定义。

定义 2.5.5 设 $(X_n,n\in \boldsymbol{N})$ 是滤波空间 $(\Omega,\mathscr{F},\mathscr{F}_n,P)$ 上的一个适应过程，若

$$E|X_n|<\infty,\quad n\in \boldsymbol{N} \tag{2.5.7}$$

且

$$E(X_{n+1}|\mathscr{F}_n)=X_n,\text{a.s.} \tag{2.5.8}$$

则称 $(X_n,n\in \boldsymbol{N})$ 为离散时间鞅或简称为离散鞅。

若式 (2.5.8) 改为

$$E(X_{n+1}|\mathscr{F}_n)\geqslant X_n,\ \text{a.s.} \tag{2.5.9}$$

则 $(X_n,n\in \boldsymbol{N})$ 称为下鞅。

若 $(-X_n,n\in \boldsymbol{N})$ 是一个下鞅，那么 $(X_n,n\in \boldsymbol{N})$ 称为上鞅。对于上鞅有

$$E(X_{n+1}|\mathscr{F}_n)\leqslant X_n,\ \text{a.s} \tag{2.5.10}$$

对于连续时间的情形可以作如下的定义：

定义 2.5.6 设 $\{X_t\}_{t\in R_+}$是$(\Omega,\mathscr{F},\mathscr{F}_t,P)$ 上的随机变量族，如果 $E|X_t|<\infty$, $t\in R_+$, X_t 关于 $\mathscr{F}_t$ 可测且

$$E(X_t|\mathscr{F}_s)=X_s,\ \text{a.s.},\forall t>s \tag{2.5.11}$$

那么 $(X_t,t\in R_+)$ 称为鞅。

若

$$E(X_t|\mathscr{F}_s)\geqslant X_s,\text{a.s.},\forall t>s \tag{2.5.12}$$

则称 $(X_t,t\in R_+)$ 为下鞅。

若

$$E(X_t|\mathscr{F}_s)\leqslant X_s,\ \text{a.s.},\forall t>s \tag{2.5.13}$$

则称 $(X_t,t\in R_+)$ 为上鞅。

若$(X_t,t\in R_+)$为下鞅，即式(2.5.12)成立. 在式(2.5.12)两边取数学期望得

$$E(E(X_t|\mathscr{F}_s))=EX_t\geqslant EX_s,\forall t>s$$

同理，对于鞅有

$$E(E(X_t|\mathscr{F}_s))=EX_t=EX_s,\forall t>s$$

对于上鞅有

$$E(E(X_t|\mathscr{F}_s))=EX_t\leqslant EX_s,\forall t>s$$

总而言之，上鞅的数学期望不增加；下鞅的数学期望不减小；鞅的数学期望等于常数。

例1 设 $\{\xi_n,n\in N\}$ 是$(\Omega,\mathscr{F},\mathscr{F}_n,P)$上的随机变量序列，$\xi_n$ 在条件 $\sigma(\xi_0,\cdots,\xi_{n-1})$ 下的条件数学期望为零. 若 $X_n=\sum\limits_{k=0}^{n}\xi_k$, 那么$(X_n,n\in N)$是鞅，这里 $\mathscr{F}_n=\sigma(\xi_0,\cdots,\xi_n)$.

解

$$\begin{aligned}E(X_{n+1}|\mathscr{F}_n)&=E(X_n+\xi_{n+1}|\mathscr{F}_n)\\&=E(X_n|\mathscr{F}_n)+E(\xi_{n+1}|\mathscr{F}_n)\end{aligned}$$

因为 $E(\xi_{n+1}|\mathscr{F}_n)=0$, X_n 关于 $\mathscr{F}_n$ 可测，由引理 2.4.4

得

$$E(X_n|\mathscr{F}_n)=X_n, \text{ a.s.}$$

因此有

$$E(X_{n+1}|\mathscr{F}_n)=X_n \text{ a.s.}$$

所以 $(X_n, n\in \boldsymbol{N})$ 是鞅。

例 2 若 $\{\xi_n\}_{n\geq 0}$ 是均值为 1 的独立随机变量序列，即 $E\xi_n=1, n\geq 0$。设

$$X_n=\prod_{k=0}^{n}\xi_k, \quad \mathscr{F}_n=\sigma(\xi_0,\cdots,\xi_n)$$

那么 $(X_n, n\in \boldsymbol{N})$ 在滤波空间 $(\Omega,\mathscr{F},\mathscr{F}_n,P)$ 上是鞅。

验证留作习题。

例 3 设 ξ 是定义在 $(\Omega,\mathscr{F},P)$ 上的随机变量，$E|\xi|<\infty$ 且

$$\mathscr{F}_0\subset\mathscr{F}_1\subset\mathscr{F}_2\subset\cdots\subset\mathscr{F}$$

令 $X_n=E(\xi|\mathscr{F}_n)$，那么 $(X_n, n\in \boldsymbol{N})$ 是鞅。

验证留作习题。

例 4 若 $(\xi_n, n\in \boldsymbol{N})$ 是非负随机变量序列，那么 $X_n=\xi_0+\cdots+\xi_n$ 是下鞅。

事实上，令 $\mathscr{F}_n=\sigma(\xi_0,\cdots,\xi_n)$，于是

$E(X_{n+1}|\mathscr{F}_n)=E(X_n+\xi_{n+1}|\mathscr{F}_n)\geq X_n$, a.s.

例 5 若 $(X_n, n\in \mathrm{N})$ 是鞅，$g(\cdot)$ 是凸函数，$E|g(X_n)|<\infty$，那么 $(g(X_n), n\in \boldsymbol{N})$ 是下鞅。

解 利用 Jensen 不等式:

$$E(g(X_{n+1})|\mathscr{F}_n)\geq g(E(X_{n+1}|\mathscr{F}_n))=g(X_n)$$

因此 $(g(X_n), n\in \boldsymbol{N})$ 为下鞅。

最后我们用鞅的等价定义来结束本节。

定理 2.5.1 设 $(\Omega,\mathscr{F},\mathscr{F}_n,P)$ 为一滤波空间，$(X_n|n\in \boldsymbol{N})$ 为上鞅的充分必要条件为：对任意两个有界停时

$s \leqslant \tau$，有

$$E(X_\tau|\mathscr{F}_s) \leqslant X_s, \text{ a.s.} \tag{2.5.14}$$

证明 充分性： 由于式(2.5.14)成立，$n+1$和n均可看作停时，故

$$E(X_{n+1}|\mathscr{F}_n) \leqslant X_n, \text{ a.s.}$$

即 $(X_n, n \in \boldsymbol{N})$ 为上鞅。

必要性：设 $\tau \leqslant n$，于是

$$|X_\tau| \leqslant \sum_{i=0}^{n} |X_i|, \quad |X_s| \leqslant \sum_{j=0}^{n} |X_j|$$

因为 $|X_n|$ 可积，$n = 1, 2, \cdots$，所以$|X_\tau|$和$|X_s|$均可积。

设 $A \in \mathscr{F}_s$，$j \in \boldsymbol{N}$，于是有

$$A \cap [s = j] \cap [\tau > j] \in \mathscr{F}_j$$

首先假定 $\tau - s \leqslant 1$.因为上鞅的数学期望不增加，所以我们有

$$\int_A (X_s - X_\tau) dP = \sum_{j=0}^{n} \int_{A\cap[s=j]\cap[\tau>j]} (X_j - X_{j+1}) dP \geqslant 0$$

对一般情形，令

$$r_j = \tau \wedge (s + j), \quad j = 1, 2, \cdots, n$$

由引理 2.5.2，每一个 r_j 为停时，且 $s \leqslant r_1 \leqslant \cdots \leqslant r_n = \tau$

$$r_1 - s \leqslant 1, r_{j+1} - r_j \leqslant 1, \forall 1 \leqslant j \leqslant n-1$$

令 $A \in \mathscr{F}_s$，则 $A \in \mathscr{F}_{r_j}$，$1 \leqslant j \leqslant n$，利用上面的结论得

$$\int_A X_s dP \geqslant \int_A X_{r_1} dP \geqslant \cdots \geqslant \int_A X_\tau dP \tag{2.5.15}$$

由于 X_s 关于 $\mathscr{F}_s$ 可测，故

$$X_s = E(X_s|\mathscr{F}_s)$$

又根据条件数学期望的定义

$$\int_A X_s dP = \int_A E(X_s|\mathscr{F}_s)dP, \forall A \in \mathscr{F}_s$$
$$\int_A X_\tau dP = \int_A E(X_\tau|\mathscr{F}_s)dP, \forall A \in \mathscr{F}_s \tag{2.5.16}$$

由式 (2.5.15) 和 (2.5.16) 可得

$$X_s = E(X_s|\mathscr{F}_s) \geqslant E(X_\tau|\mathscr{F}_s), \text{ a.s.}$$ ▮

推论 2.5.1 设 $(\Omega, \mathscr{F}, \mathscr{F}_n, P)$ 为滤波空间，$(X_n, n \in N)$ 为鞅的充分必要条件是：对任意有界停时 $s \leqslant \tau$，有

$$E(X_\tau|\mathscr{F}_s) = X_s \tag{2.5.17}$$

证明 完全类似于定理 2.5.1。 ▮

推论 2.5.1 给出了鞅的等价定义：

定义 2.5.7 设 $(\Omega, \mathscr{F}, \mathscr{F}_n, P)$为滤波空间，$(X_n, n \in N)$为适应序列，若对任何有界停时 $s \leqslant \tau$ 有

$$E(X_\tau|\mathscr{F}_s) = X_s$$

则称 $(X_n, n \in N)$ 为鞅.

定理 2.5.2 设 $(X_n, n \in N)$ 为上鞅，τ 为停时，则

$$E|X_{\tau\wedge k}| \leqslant EX_0 + 2EX_k^- \tag{2.5.18}$$
$$E|X_\tau \chi_{[\tau<\infty]}| \leqslant 3\sup_n E|X_n| \tag{2.5.19}$$

证明 由于 $(X_n, n \in N)$为上鞅，可知 X_n^- 是下鞅(见习题 2.5,5)。

根据定理 2.5.1

$$E|X_{\tau\wedge k}| = EX_{\tau\wedge k} + 2EX_{\tau\wedge k}^- \leqslant EX_0 + 2EX_k^- \tag{2.5.20}$$

这是因为 $(X_n, n \in N)$ 为上鞅，因此有

$$EX_{k\wedge\tau} = E[E(X_{\tau\wedge k}|\mathscr{F}_0)] \leqslant EX_0$$

$(X_n^-, n \in N)$ 为下鞅，因此有

$$EX_k^- = E[E(X_k^-|\mathscr{F}_{\tau\wedge k})] \geqslant EX_{\tau\wedge k}^-$$

故式 (2.5.20) 成立。

由此可得

$$E|X_{\tau\wedge k}\chi_{[\tau<\infty]}| \leqslant EX_0 + 2EX_k^- \leqslant 3\sup_n E|X_n|$$

令 $k\to\infty$，由 Fatou 引理式(2.5.18)成立。 ∎

习　题　2.5

1. 设 $\{y_n\}$ 为一列独立正态 $N(0,1)$ 随机变量，令

$$x_n = \exp\left(y_1 + y_2 + \cdots + y_n - \frac{n}{2}\right)$$

证明 x_n 自身为鞅。

2. 设 $(X_n, n\in N)$ 是独立随机序列，$EX_n = 0$，$EX_n^2 = \sigma_n^2 < \infty$，$\forall n = 1, 2, \cdots$。令

$$Y_n = \left(\sum_{j=1}^{n} X_j\right)^2 - \sum_{i=1}^{n}\sigma_i^2$$

证明 $(Y_n, n\in N)$ 是鞅。

3. 设 $(Y_n, n\in N)$ 为上鞅，随机序列 $(Z_n, n\in N)$ 只取 0，1 两个值。令

$$X_n = Y_1 + Z_1(Y_2 - Y_1) + \cdots + Z_{n-1}(Y_n - Y_{n-1})$$

证明 $(X_n, n\in N)$ 是上鞅。

4. 证明在离散时间情况下，停时的定义等价于：τ 可以取值 $+\infty$ 并且 $\{\omega: \tau(\omega) = n\}\in\mathscr{F}_n$。

5. 若 $(X_n, n\in N)$ 为上鞅，证明 $(X_n^-, n\in N)$ 为下鞅。

6. 若 τ 为停时，c 为常数，证明 $\tau + c$ 也是停时。

7. 验证例2。

8. 验证例3。

2.6 鞅与上鞅的不等式

提要　上鞅不等式是鞅收敛定理证明的依据。本节证明了 Kolmogorov 不等式，Doob 不等式和著名的上穿不等式。

鞅和上鞅的不等式是下面证明鞅收敛定理的工具。这里证明几个著名的上鞅不等式。

定理 2.6.1 设 $(X_n, n\in \boldsymbol{N})$ 为一上鞅，那么对任何 $\lambda > 0$，我们有

$$(1)\quad \lambda P\,[\sup_{n\leqslant k} X_n \geqslant \lambda] \leqslant EX_0 - \int_{[\sup\limits_{n\leqslant k} X_n<\lambda]} X_k dP;\quad (2.6.1)$$

$$(2)\quad \lambda P\,[\inf_{n\leqslant k} X_n \leqslant -\lambda] \leqslant \int_{[\inf\limits_{n\leqslant k} X_n\leqslant-\lambda]} (-X_k) dP;\quad (2.6.2)$$

$$(3)\quad \lambda P\,[\sup_{n\leqslant k} |X_n| \geqslant \lambda] \leqslant EX_0 + 2EX_k^-. \quad (2.6.3)$$

证明 令 $\tau = \inf\{n\in \boldsymbol{N}: X_n \geqslant \lambda\}\wedge k$，由于初遇时间是停时，由引理 2.5.2，$\tau$ 是有界停时：

$$\text{在集 } [\sup_{n\leqslant k} X_n \geqslant \lambda] \text{ 上，} X_\tau \geqslant \lambda$$

$$\text{在集 } [\sup_{n\leqslant k} X_n < \lambda] \text{ 上，} \tau = k$$

根据定理 2.5.1，$(X_n, n\in \boldsymbol{N})$ 为上鞅，因此

$$X_s \geqslant E(X_\tau|\mathscr{F}_s), \forall s\leqslant \tau$$

$$EX_s \geqslant E[E(X_\tau|\mathscr{F}_s)] = EX_\tau$$

又

$$EX_0 \geqslant EX_\tau = \int_{[\sup\limits_{n\leqslant k} X_n\geqslant\lambda]} X_\tau dP + \int_{[\sup\limits_{n\leqslant k} X_n<\lambda]} X_\tau dP$$

$$\geqslant \lambda P[\sup_{n\leqslant k} X_n \geqslant \lambda] + \int_{[\sup\limits_{n\leqslant k} X_n<\lambda]} X_k dP$$

故

$$\lambda P\,[\sup_{n\leqslant k} X_n \geqslant \lambda] \leqslant EX_0 - \int_{[\sup\limits_{n\leqslant k} X_n<\lambda]} X_k dP$$

成立。

同理可证式 (2.6.2)。

令 $s = \inf\{n\in \boldsymbol{N}: X_n \leqslant -\lambda\}\wedge k$，$s$ 是有界停时：

$$在集\ \{\inf_{n\leqslant k} X_n \leqslant -\lambda\}\ 上，X_s \leqslant -\lambda$$

$$在集\ \{\inf_{n\leqslant k} X_n > -\lambda\}\ 上，s = k$$

又
$$EX_k \leqslant EX_s = \int_{[\inf\limits_{n\leqslant k} X_n \leqslant -\lambda]} X_s dP + \int_{[\inf\limits_{n\leqslant k} X_n > -\lambda]} X_s dP$$

$$\leqslant -\lambda P[\inf_{n\leqslant k} X_n \leqslant -\lambda] + \int_{[\inf\limits_{n\leqslant k} X_n > -\lambda]} X_k dP$$

移项可得

$$\lambda P[\inf_{n\leqslant k} X_n \leqslant -\lambda] \leqslant \int_{[\inf\limits_{n\leqslant k} X_n > -\lambda]} X_k dP - \int_{\Omega} X_k dP$$

即

$$\lambda P[\inf_{n\leqslant k} X_n \leqslant -\lambda] \leqslant -\int_{[\inf\limits_{n\leqslant k} X_n \leqslant -\lambda]} X_k dP$$

由式 (2.6.1) 和 (2.6.2) 得

$$\lambda P[\sup_{n\leqslant k}|X_n| \geqslant \lambda] \leqslant EX_0 + 2EX_k^- \leqslant 3\sup_{n\leqslant k} E|X_n| \quad ▮$$

推论 2.6.1（Kolmogorov 不等式） 设 $(X_n,\ n\in \boldsymbol{N})$ 是鞅，$EX_k^2 < \infty$，于是对任何 $\lambda > 0$，有

$$P[\sup_{n\leqslant k}|X_n| \geqslant \lambda] \leqslant \frac{1}{\lambda^2} EX_k^2 \tag{2.6.4}$$

证明 由 X_n 是鞅，有

$$X_n = E(X_{n+1}|\mathscr{F}_n)$$

利用 Jensen 不等式

$$X_n^2 = (E(X_{n+1}|\mathscr{F}_n))^2 \leqslant E(X_{n+1}^2|\mathscr{F}_n)$$

所以 X_n^2 是下鞅，$(-X_n^2, \mathscr{F}_n)$ 是上鞅。用定理 2.6.1 中的式 (2.6.2) 可得

$$\lambda^2 P[\sup_{n\leqslant k}|X_n| \geqslant \lambda] = \lambda^2 P[\inf_{n\leqslant k}(-X_n^2) \leqslant -\lambda^2] \leqslant E|X_k|^2$$

由此得

$$P[\sup_{n\leqslant k}|X_n| \geqslant \lambda] \leqslant \frac{E|X_k|^2}{\lambda^2} \quad ▮$$

为了证明上穿不等式，需要引进一些记号：设 $(X_n, \mathscr{F}_n)$ 为随机序列，$[a,b]$ 为有限区间，令

$$
\begin{aligned}
&\tau_0 = \inf\{n: X_n < a\} \text{（初遇时间）}\\
&\tau_1 = \inf\{n: n > \tau_0,\ X_n > b\}\\
&\cdots\\
&\tau_{2j} = \inf\{n: n > \tau_{2j-1}, X_n < a\}\\
&\tau_{2j+1} = \inf\{n: n > \tau_{2j}, X_n > b\}\\
&\cdots
\end{aligned}
$$

可以证明 $\{\tau_k\}$ 是一列上升的停时。事实上，τ_0 为集 $\{X_n < a\}$ 的初遇时间(由定义 2.5.2 可知，它是停时)，τ_1 为集 $\{\omega: X_n > b\} \cap (\tau_0, \infty)$ 的初遇时间，……。显然 $\tau_0 \leqslant \tau_1 \leqslant \tau_2 \leqslant \cdots$。

X_n 在下标 n 从 1 变到 N 的过程中，从闭区间 $[a,b]$ 的左边穿越到 $[a,b]$ 右边的次数(返回不算次数)称为 $\{X_n\}_{n=1,\cdots,N}$ 对于闭区间 $[a, b]$ 的上穿次数，记作 $V_a^b(X, N)$，于是 $V_a^b(X,k)$ 是 $\mathscr{F}_k$ 可测随机变量。事实上

$$
\begin{aligned}
&[V_a^b(X,k) = j] = [\tau_{2j-1} \leqslant k < \tau_{2j+1}]\\
&= [\tau_{2j-1} \leqslant k] \cap [k < \tau_{2j+1}] \in \mathscr{F}_k
\end{aligned}
\tag{2.6.5}
$$

定理 2.6.2（**上穿不等式定理**） 设 $(X_n, \mathscr{F}_n)$ 为上鞅，$N \in \boldsymbol{N}$，于是

$$
E\{V_a^b(X, N)\} \leqslant \frac{E(X_N - a)^-}{b - a}
\tag{2.6.6}
$$

证明 我们继续采用上面的记号。由停时 τ_n 的定义可知，$X_{\tau_{2j}} < a$，$X_{\tau_{2j+1}} > b$，因此

$$
X_{\tau_{2j+1}} - X_{\tau_{2j}} > b - a
\tag{2.6.7}
$$

定义

$$
v_k = \begin{cases} 0, & k < \tau_0 \\ 1, & \tau_{2j} \leqslant k < \tau_{2j+1}, \forall j = 0, 1, \cdots \\ 0, & \tau_{2j+1} \leqslant k < \tau_{2j+2} \end{cases}
$$

ν_k 是 $\mathscr{F}_k$ 可测的随机变量. 由式 (2.6.5) 得

$$\{V_a^b(X,N)=j\}=\{\tau_{2j-1}\leqslant N<\tau_{2j}\}\cup\{\tau_{2j}\leqslant N<\tau_{2j+1}\}$$

在集 $\{\tau_{2j-1}\leqslant N<\tau_{2j}\}$ 上,由式 (2.6.7) 有

$$\sum_{k=1}^{N-1}\nu_k(X_{k+1}-X_k)=\sum_{k=1}^{[\frac{N}{2}]}(X_{\tau_{2k-1}}-X_{\tau_{2k-2}})$$

$$\geqslant(b-a)V_a^b(X,N)$$

在集 $\{\tau_{2j}\leqslant N<\tau_{2j+1}\}$ 上

$$\sum_{k=1}^{N-1}\nu_k(X_{k+1}-X_k)=\sum_{k=1}^{j}(X_{\tau_{2j-1}}-X_{\tau_{2j-2}})$$

$$+(X_N-X_{\tau_{2j}})\geqslant V_a^b(X,N)(b-a)-(X_N-a)^-$$

总之,我们有

$$\sum_{k=1}^{N-1}\nu_k(X_{k+1}-X_k)\geqslant(b-a)V_a^b(X,N)-(X_N-a)^- \tag{2.6.8}$$

由于 $\{\nu_k=1\}\in\mathscr{F}_k$, 由上鞅的性质可得

$$E\left\{\sum_{k=1}^{N-1}\nu_k(X_{k+1}-X_k)\right\}=\sum_{k=1}^{N-1}\int_{\{\nu_k=1\}}(X_{k+1}-X_k)dP\leqslant 0$$

(见习题 2.6,1). 由式 (2.6.8) 得

$$(b-a)E\{V_a^b(X,N)\}-E(X_N-a)^-\leqslant 0$$

即式 (2.6.6) 成立. ∎

定理 2.6.3 (Doob 不等式定理) 设 $(X_n, n\in\boldsymbol{N})$ 是一非负下鞅,令

$$X^*=\sup_n X_n$$

那么有

$$EX^*\leqslant\frac{e}{e-1}(1+\sup_n E(X_n\log^+X_n)) \tag{2.6.9}$$

$$\|X^*\|_p\leqslant q\sup_n\|X_n\|_p \tag{2.6.10}$$

其中 $p>1$, $q>1$ 为一对共轭指数,即

$$\frac{1}{p}+\frac{1}{q}=1,\ \|\xi\|_p=(E|\xi|^p)^{1/p}$$

$\log^+ a$ 表示 $\log a$ 的正部。

证明 设 $k\in N$, 令 $X_k^*=\sup\limits_{n\leqslant k}X_n$. 令

$$F(\lambda)=P(X_k^*>\lambda),\ 对于\ \lambda>0$$

由于

$$[\inf_{n\leqslant k}(-X_n)\leqslant-\lambda]=[\sup_{n\leqslant k}X_n\geqslant\lambda]$$

根据式 (2.6.2) 可得

$$\lambda P[X_k^*\geqslant\lambda]\leqslant\int_{[X_k^*\geqslant\lambda]}X_k dP$$

因此有

$$F(\lambda)\leqslant\frac{1}{\lambda}\int_{[X_k^*>\lambda]}X_k dP$$

设 $\Psi(\lambda)$ 为 $[0,\infty)$ 上的右连续函数且 $\Psi(0)=0$. 而且设 $F_{X^*}(\lambda)$ 是 X_k^* 的分布函数,即

$$F_{X^*}(\lambda)=P(X_k^*\leqslant\lambda)=1-F(\lambda)$$

于是

$$E[\Psi(X_k^*)]=\int_0^\infty\Psi(\lambda)dF_{X^*}(\lambda)=-\int_0^\infty\Psi(\lambda)dF(\lambda)$$

由分部积分法得

$$\begin{aligned}
-\int_0^\infty\Psi(\lambda)dF(\lambda)&=\int_0^\infty F(\lambda)d\Psi(\lambda)-\Psi(\lambda)F(\lambda)\Big|_0^\infty\\
&\leqslant\int_0^\infty F(\lambda)d\Psi(\lambda)\\
&\leqslant\int_0^\infty\left(\frac{1}{\lambda}\int_{[X_k^*>\lambda]}X_k dP\right)d\Psi(\lambda)\\
&=E\left[X_k\left(\int_0^{X_k^*}\frac{d\Psi(\lambda)}{\lambda}\right)\right]
\end{aligned}$$

这最后一个等式利用了 Fubini 定理。

如果令 $\Psi(\lambda)=(\lambda-1)^+$，那么有

$$E(X_k^*-1)\leqslant E(X_k^*-1)^+\leqslant EX_k\log^+X_k^* \quad (2.6.11)$$

由于 $\log x\leqslant x/e$ $(x\geqslant 0)$，故对任意 $a\geqslant 0, b\geqslant 0$（取 $x=b/a$）有

$$a\log^+b\leqslant a\log^+a+b/e$$

因而

$$E[X_k\log^+X_k^*]\leqslant E[X_k\log^+X_k]+\frac{1}{e}E[X_k^*]$$

于是从式 (2.6.11) 得

$$EX_k^*\leqslant\frac{e}{e-1}(1+E[X_k\log^+X_k])$$

由于 $X_k^*\uparrow X^*$，在上面的不等式中令 $k\to\infty$，利用 Fatou 引理得

$$EX^*\leqslant\frac{e}{e-1}(1+\sup_n E(X_n\log^+X_n))$$

现在证明式 (2.6.10) 成立。

如果令 $\Psi(\lambda)=\lambda^p$，$\lambda>0$，那么有

$$E(X_k^{*p})\leqslant\frac{p}{p-1}E(X_kX_k^{*p-1})=qE(X_kX_k^{*p-1})$$

再利用 Hölder 不等式(定理 1.6.2) 得

$$E(X_k^{*p})\leqslant q(EX_k^p)^{1/p}(EX_k^{*p})^{1/q} \quad (2.6.12)$$

为了证明式 (2.6.10)，不妨设 $\sup_n\|X_n\|_p<\infty$，于是有

$$\|X_k^*\|_p\leqslant\left\|\sum_{n=0}^{k}X_n\right\|_p\leqslant\sum_{n=0}^{k}\|X_n\|_p<\infty$$

在式 (2.6.12) 两边用 $(EX_k^{*p})^{1/q}$ 去除，可得

$$\|X_k^*\|_p\leqslant q\|X_k\|_p$$

利用 Fatou 引理可得式 (2.6.10)。 ∎

习 题 2.6

1.设 $(X_n, n\in N)$ 是上鞅，$A\in\mathscr{F}_k$，证明

$$\int_A (X_{k+1}-X_k)dP\leqslant 0$$

2.设 $(X_n, n\in N)$ 为上鞅，证明对 $\lambda>0$

$$\tau=\inf\{n\in N: X_n\leqslant-\lambda\}\wedge k$$

是有界停时.

3.设 $(X_n, n\in N)$ 为上鞅，证明

$$\sup_n E|X_n|\leqslant EX_0+2\sup_n EX_n^-$$

2.7 鞅的收敛定理

提要 本节讨论上鞅的收敛定理，一致可积上鞅的收敛定理。定理证明的依据是上穿不等式及一致可积性等概念。

上鞅收敛定理是鞅极限定理的基础，以后的许多定理都以此为出发点。

定理 2.7.1 设 $(X_n, n\in N)$ 为上鞅，若 $\sup_n EX_n^-<\infty$，那么当 $n\to\infty$ 时，X_n 几乎处处收敛于一个可积随机变量 X_∞。若 $(X_n, n\in N)$ 为非负上鞅，则 $n\in N$ 时

$$E(X_\infty|\mathscr{F}_n)\leqslant X_n,\ \text{a.s.} \tag{2.7.1}$$

证明 设 Q 为有理数全体。对固定的 $a, b\in Q$，$a<b$，上穿次数 $V_a^b(X, N)$ 是 N 的单调上升函数，令

$$\lim_{N\to\infty}V_a^b(X,N)=V_a^b(X)$$

表示序列 $\{X_n\}_{n\geqslant1}$ 上穿 $[a,b]$ 的次数。

利用定理 2.6.2 中的上穿不等式 (2.6.6)，可得

$$E\{V_a^b(X)\} = \lim_{N\to\infty} E\{V_a^b(X,N)\}$$
$$\leqslant \frac{1}{b-a} \sup_n E(X_n - a)^-$$
$$\leqslant \frac{|a| + \sup_n EX_n^-}{b-a} < \infty$$

由于 $V_a^b(X) < \infty$, a.s.,于是
$$P[V_a^b(X) < \infty] = 1$$
记
$$W_{a,b} = \{\liminf_n X_n < a < b < \limsup_n X_n\}$$
$$W = \bigcup_{\substack{a<b \\ a,b\in Q}} W_{a,b} = \{\liminf_n X_n \neq \limsup_n X_n\}$$
容易看到,在 $W_{a,b}$ 上 $V_a^b(X) = +\infty$, 因此
$$P(W_{a,b}) = 0$$
因而 $P(W) = 0$, 即
$$\lim_n X_n = X_\infty, \text{ a.s.}$$

由 Fatou 引理
$$E|X_\infty| \leqslant \sup_n E|X_n| \leqslant EX_0 + 2\sup_n EX_n^- < \infty$$
这后面的不等式是根据上鞅 $EX_0 \geqslant EX_n$ 获得的。因此 X_∞ 可积。

如果 $\{X_n\}$ 非负,则对任何 $m > n$ 有
$$E[X_m|\mathscr{F}_n] \leqslant X_n, \text{ a.s.}$$
令 $m\to\infty$, 由 Fatou 引理可得
$$E[X_\infty|\mathscr{F}_n] \leqslant X_n, \text{ a.s.}$$ ∎

定理 2.7.2 设 $(X_n, n\in \boldsymbol{N})$ 为上鞅,如果 $\{X_n\}$ 一致可积,则存在一个可积随机变量 X_∞, 使得 $X_n \xrightarrow{\text{a.s.}, L_1} X_\infty$ 并且对一切 $n\in \boldsymbol{N}$, 有

$$E[X_\infty|\mathscr{F}_n] \leqslant X_n, \text{ a.s.}$$

证明 由于 $\{X_n\}$ 一致可积,有

$$\int |X_n| dP \leqslant c + \int_{[|X_n| \geqslant c]} |X_n| dP$$

其中

$$\int_{[|X_n| \geqslant c]} |X_n| dP < \varepsilon, \ \forall c > c_0$$

因此

$$\sup_n E|X_n| < \infty$$

由定理 2.7.1, $X_n \to X_\infty$, a.s.

下面证明 $X_n \xrightarrow{L_1} X_\infty$.

根据定理 2.4.2, $\lim_n X_n = X_\infty$, a.s., X_∞ 可积, $\{X_n\}$ 一致可积,因此

$$X_n \xrightarrow{L_1} X_\infty$$

由于$(X_n, n \in N)$为上鞅,对 $m > n$, $E(X_m|\mathscr{F}_n) \leqslant X_n$ 而

$$\begin{aligned} & E|E(X_m|\mathscr{F}_n) - E(X_\infty|\mathscr{F}_n)| \\ \leqslant & E(E(|X_m - X_\infty||\mathscr{F}_n)) \\ = & E|X_m - X_\infty| \xrightarrow[m\to\infty]{} 0 \end{aligned}$$

故得

$$E(X_\infty|\mathscr{F}_n) \leqslant X_n, \text{ a.s.} \qquad \blacksquare$$

推论 2.7.1 在上面的定理中如果 $(X_n, \mathscr{F}_n)$ 为鞅,则

$$E(X_\infty|\mathscr{F}_n) = X_n, \text{ a.s.}$$

证明 由于 $(X_n, n \in N)$为鞅,因此 $(X_n, n \in N)$ 也是上鞅,故由上面的定理得

$$E(X_\infty|\mathscr{F}_n) \leqslant X_n$$

由于 $(X_n, n \in N)$ 也是下鞅,故

$$E(-X_\infty|\mathscr{F}_n) \leqslant -X_n$$

即

$$E(X_\infty|\mathscr{F}_n) \geqslant X_n$$

于是有

$$E(X_\infty|\mathscr{F}_n) = X_n, \text{ a.s.}$$ ∎

推论 2.7.2 设 $(X_n, n\in N)$是上鞅，$EX_n^2 < c$，$\forall n\in N$，c 为常数，那么 $X_n \to X_\infty$，a.s. 且 X_∞ 可积。

证明 由 Jensen 不等式，有

$$(E|X_n|)^2 \leqslant EX_n^2 < c$$

故

$$\sup_n EX_n^- \leqslant \sup_n E|X_n| \leqslant c$$

满足定理 2.7.1，因此结论成立。 ∎

非空集族 $\mathscr{C}$ 称为 π 族，如果它满足若 A，$B\in\mathscr{C}$，则 $A\cap B\in\mathscr{C}$。

集族 λ 称为 λ 族，如果它满足下列条件：

(1) $\Omega\in\lambda$；

(2) 若 $A,B\in\lambda$，且 $A\supset B$ 则 $A-B\in\lambda$；

(3) 设 $\{E_n\}$ 为 λ 中的增序列，则

$$\lim_n E_n \in \lambda$$

显然由 Ω 一切子集所成的族是 λ 族。

设 $\mathscr{C}$ 是 π 族，$\lambda(\mathscr{C})$ 和 $\sigma(\mathscr{C})$ 分别表示由 $\mathscr{C}$ 产生的 λ 族和 σ 代数，则 $\lambda(\mathscr{C}) = \sigma(\mathscr{C})$[3]。

利用这一命题可以证明：

引理 2.7.1 设 $(\Omega,\mathscr{F},P)$ 为概率空间，ξ，η 为两个可积随机变量，而且设 $\mathscr{C}\subset\mathscr{F}$，$\mathscr{C}$ 是 π 族。如果 $E\xi = E\eta$ 且

$$E(\xi\chi_A) = E(\eta\chi_A),\ \forall A\in\mathscr{C}$$

那么

$$E(\xi|\sigma(\mathscr{C}))=E(\eta|\sigma(\mathscr{C}))$$

证明 留作习题。

定理 2.7.3 设 ξ 为一可积随机变量，令 $\xi_n=E(\xi|\mathscr{F}_n)$，$\eta=E(\xi|\bigvee_n \mathscr{F}_n)$，其中 $\bigvee_n \mathscr{F}_n$ 表示包含所有 $\{\mathscr{F}_n\}$ 的最小 σ 代数，于是当 $n\to\infty$ 时

$$\xi_n \xrightarrow{\text{a.s.},L_1} \eta$$

证明 因为

$$\begin{aligned}E[\xi_{n+1}|\mathscr{F}_n]&=E(E[\xi|\mathscr{F}_{n+1}]|\mathscr{F}_n)\\&=E[\xi|\mathscr{F}_n]=\xi_n\end{aligned}$$

故 $(\xi_n,\mathscr{F}_n)$ 为鞅。

再证明 $\{\xi_n\}$ 一致可积。令 $Y_i=E[|\xi||\mathscr{F}_i]$，对任何

$c>0$，我们有

$$P[Y_i\geqslant c]\leqslant\frac{1}{c}EY_i=\frac{1}{c}E|\xi|$$

于是

$$\int_{[Y_i\geqslant c]}Y_i dP=\int_{[Y_i\geqslant c]}|\xi|dP\leqslant\delta P[Y_i\geqslant c]$$

$$+\int_{[|\xi|\geqslant\delta]}|\xi|dP\leqslant\frac{\delta}{c}E|\xi|+\int_{[|\xi|\geqslant\delta]}|\xi|dP$$

给定 $\varepsilon>0$，取 δ 使

$$\int_{[|\xi|\geqslant\delta]}|\xi|dP\leqslant\frac{\varepsilon}{2}$$

于是当 $c\geqslant\frac{2\delta}{\varepsilon}E|\xi|$ 时

$$\int_{[Y_i \geq c]} Y_i dP \leqslant \varepsilon$$

这表明 $\{Y_i\}$ 是一致可积的，但是

$$Y_i = E[|\xi| | \mathscr{F}_i] \geqslant |E[\xi | \mathscr{F}_i]| = |\xi_i|$$

因而 $\{\xi_i\}$ 为一致可积族。

由定理 2.7.2，$\xi_n \xrightarrow{a.s., L_1} \xi_\infty$。设 $A \in \bigcup_n \mathscr{F}_n$，则存在 n，$A \in \mathscr{F}_n$，于是

$$E[\xi_\infty \chi_A] = E[\xi_n \chi_A] = E[\xi \chi_A] = E[\eta \chi_A]$$

由于 ξ_∞，η 为 $\bigvee_n \mathscr{F}_n$ 可测，而 $\bigvee_n \mathscr{F}_n = \sigma\left(\bigcup_n \mathscr{F}_n\right)$。故由引理 2.7.1 有

$$\xi_\infty = E\left[\xi_\infty \middle| \bigvee_n \mathscr{F}_n\right] = E\left[\eta \middle| \bigvee_n \mathscr{F}_n\right] = \eta \qquad \blacksquare$$

现在我们考虑以 $Z_- = \{\cdots, -2, -1, 0\}$ 为参数集的上鞅收敛定理。设 $(\Omega, \mathscr{F}, P)$ 为概率空间，$\{\mathscr{F}_n\}_{n \in Z_-}$ 为一列 $\mathscr{F}$ 的子 σ 代数。当 $n, m \in Z_-$ 时，如果 $n > m$，$\mathscr{F}_m \subset \mathscr{F}_n$，$X_n$ 关于 $\mathscr{F}_n$ 可测，$\forall n \in Z_-$，那么 $(X_n, n \in Z_-)$ 称为上鞅，如果 $\forall n \in Z_-$，X_n 可积，则有

$$E[X_n | \mathscr{F}_{n-1}] \leqslant X_{n-1}, \text{ a.s.}$$

定理2.7.4 设 $(X_n, n \in Z_-)$ 是上鞅，如果

$$\lim_{n \to -\infty} E[X_n] < \infty$$

则 $\{X_n\}$ 一致可积，且

$$X_n \xrightarrow{a.s., L_1} X_{-\infty}, \text{ 当 } n \to -\infty$$

证明 设 $V_a^b(X, -k)$ 表示序列 $\{X_{-k}, X_{-k+1}, \cdots, X_0\}$ 上穿区间 $[a, b]$ 的次数。由定理 2.6.2 知

$$EV_a^b[X, -k] \leqslant \frac{1}{b-a} E[(X_0 - a)^-]$$

令

$$V_a^b(X) = \lim_{k\to\infty} V_a^b(X,-k)$$

于是有

$$EV_a^b(X) \leqslant \frac{1}{b-a} E[(X_0 - a)^-] < \infty$$

由于 $V_a^b(X)$ 是序列 $(-X_0, -X_{-1}, \cdots)$ 上穿 $[-b, -a]$ 的次数，故由定理 2.7.1 的证明知

$$X_n \xrightarrow{\text{a.s.}} X_{-\infty}$$

当 $n \to -\infty$ 时，$EX_n \to c > -\infty$。由定理的假定知 $c < \infty$。

往证，$\{X_n\}$ 一致可积：由于 $\{E(X_0|\mathscr{F}_n)\}_{n\in Z_-}$ 一致可积(见定理 2.7.3的证明)，因此只需证明 $\{X_n - E(X_0|\mathscr{F}_n)\}$ 一致可积。因后者是非负上鞅，故我们假定 X_n 非负因而只需证明 $\{X_n\}$ 一致可积。给定 $\varepsilon > 0$，取自然数 j 足够大，从而使 $c - E[X_{-j}] < \varepsilon/2$，由上鞅性质，有

$$\begin{aligned}
\int_{[X_n>c_1]} X_n dP &= E[X_n] - \int_{[X_n\leqslant c_1]} X_n dP \\
&\leqslant EX_n - \int_{[X_n\leqslant c_1]} X_{-j} dP \\
&= EX_n - E(X_{-j}) + \int_{[X_n>c_1]} X_{-j} dP
\end{aligned}$$

对于 $c_1 > 0$，$n < -j$

由于 $c \geqslant EX_n \geqslant EX_{-j}$，故 $EX_n - EX_{-j} < \varepsilon/2 (n < -j)$。另一方面，由于

$$P[X_n > c_1] \leqslant \frac{1}{c_1} EX_n \leqslant \frac{c}{c_1}$$

故当 c_1 足够大时，对一切 $n \in Z_-$ 有

$$\int_{[X_n>c_1]} X_{-j} dP < \frac{\varepsilon}{2}$$

并且对 $i=0,-1,\cdots,-l$ 有

$$\int_{[X_i>c_1]} X_i dP < \varepsilon$$

于是，当 c_1 足够大时

$$\sup_n \int_{[X_n>c_1]} X_n dP < \varepsilon$$

这表明 $\{X_n\}$ 一致可积。由定理 2.4.2 知

$$X_n \xrightarrow{L_1} X_{-\infty}$$ ▌

推论 2.7.3 设 ξ 为可积随机变量，$\{\mathscr{G}_n\}_{n\in N}$ 为单调下降的 $\mathscr{F}$ 的子 σ 代数。令 $\xi_n=E(\xi|\mathscr{G}_n)$，则

$$\xi_n \xrightarrow{a.s.,L_1} E\left(\xi \Big| \bigcap_n \mathscr{G}_n\right)$$

证明 由上面的定理知，存在 η 使

$$\xi_n \xrightarrow{a.s.,L_1} \eta$$

设

$$A \in \bigcap_n \mathscr{G}_n$$

于是有

$$\lim_{n\to\infty} E(\xi_n \chi_A) = E(\eta \chi_A)$$

但对一切 n，$E[\xi_n \chi_A]=E[\xi \chi_A]$，由于 η 关于 $\bigcap\limits_n \mathscr{G}_n$ 可测，由条件数学期望定义得

$$\eta = E\left(\xi \Big| \bigcap_n \mathscr{G}_n\right)$$ ▌

下面的定理对系统辨识的收敛性证明是有用的。

定理 2.7.5 设 $(S_n,\mathscr{F}_n)$ 为下鞅，$ES_1^- < \infty$，$\{\mathscr{F}_n\}$ 是 $\{Z_n\}$，$\{Y_n\}$ 的适应 σ 代数子序列。令 $Z_1=Y_1=0$，有

$$S_n \leqslant Z_n + Y_n, \quad n \geqslant 2 \tag{2.7.2}$$

对于 $b > 0$，令

$$\tau = \inf\{n: Y_n > b\} \tag{2.7.3}$$

且

$$EY_\tau \chi_{[\tau<\infty]} < \infty \tag{2.7.4}$$

那么 S_n 在

$$\{\sup Z_n < \infty, \ \sup Y_n < b\} \tag{2.7.5}$$

上几乎处处收敛。

证明 对任给 $a > 0$，令

$$\tau' = \inf\{n: Z_n > a\}$$

并且设

$$\omega_n = \min(\tau' - 1, \tau, n)$$

则 ω_n 是一有界停时序列，且 $\omega_n \leqslant \omega_{n+1}$($Z_n$可见 2.9 节)。

因此，由定理 2.5.1 知 $(S_{\omega_n}, \mathscr{F}_{\omega_n})$ 也是下鞅. $ES_1^- < \infty$ 隐含着 $ES_n^- < \infty$，

$$\begin{aligned} ES_{\omega_n}^+ &\leqslant E(Z_{\omega_n}^+) + \sum_{j=1}^{n} E(Y_j^+ \chi_{[\omega_n = j]}) \\ &\leqslant a + \sum_{j=1}^{n} EY_j^+ (\chi_{[\tau'-1 > j = \tau]} + \chi_{[\tau'-1 = j < \tau]}) \\ &\quad + EY_n^+ \chi_{[\tau'-1 > n < \tau]} \\ &\leqslant a + b \sum_{j=1}^{n} EY_j^+ \chi_{[\tau'-1 > j = \tau]} \\ &\leqslant a + b + EY_\tau^+ \chi_{[\tau<\infty]} \end{aligned}$$

由条件式 (2.7.4) 可知

$$ES_{\omega_n}^+ < \infty$$

由定理 2.7.1 可得 $\lim\limits_{n} S_{\omega_n}$ 极限存在且有限。

当 $\tau = \tau' = \infty$ 时，S_n在集 $\{\sup Z_n < a, \ \sup Y_n < b\}$

上 a.s. 收敛. 由于 a 的任意性，可知 S_n 在集 (2.7.5) 上 a.s. 收敛. ▮

推论 2.7.4 条件 (2.7.4) 用下列条件 (2.7.6) 代替：

$$E(\sup_{n\geq 1}(Y_{n+1}-Y_n^+))<\infty \tag{2.7.6}$$

定理 2.7.5 仍然成立.

证明 设 τ 由式 (2.7.3) 定义

$$EY_\tau\chi_{[\tau<\infty]}=EY_{\tau-1}^+\chi_{[\tau<\infty]}+EY_\tau\chi_{[\tau<\infty]}-EY_{\tau-1}^+\chi_{[\tau<\infty]}$$
$$\leqslant E(\sup_n(Y_{n+1}-Y_n^+))+EY_{\tau-1}^+\chi_{[\tau<\infty]} \tag{2.7.7}$$

由条件 (2.7.6), $t\geqslant 0$ 和 $Y_1=0$ 可推出

$$EY_\tau\chi_{[\tau<\infty]}<\infty$$

即满足定理 2.7.5 的条件，故推论成立. ▮

推论 2.7.5 设 $(S_n,\mathscr{F}_n)$ 是下鞅，Y_n 是 $\mathscr{F}_n$ 可测的随机变量序列，$E|Y_n|<\infty$, $E(\sup(Y_{n+1}^- - Y_n^-))<\infty$ 且 $S_n\leqslant Y_n$, 那么 S_n 在集

$$\left\{\sup\sum_{n=1}^{m}(E(Y_{n+1}|\mathscr{F}_n)-Y_n)<\infty,\sup Y_n^-<\infty\right\} \tag{2.7.8}$$

上 a.s. 收敛.

证明 令

$$Z_m=\sum_{2}^{m}(E(Y_n|\mathscr{F}_{n-1})-Y_n)$$

容易验证 $(Z_m,\mathscr{F}_m)$ 是鞅：

$$Z_m=E(Y_m|\mathscr{F}_{m-1})-Y_m+E(Y_{m-1}|\mathscr{F}_{m-2})$$
$$-Y_{m-1}+\cdots+E(Y_2|\mathscr{F}_1)-Y_2$$
$$=-Y_m+\sum_{3}^{m}(E(Y_n|\mathscr{F}_{n-1})-Y_{n-1})$$
$$+E(Y_2|\mathscr{F}_1)\leqslant Y_m^-+U_m \tag{2.7.9}$$

其中

$$U_m = \sum_{3}^{m} (E(Y_n|\mathscr{F}_{n-1}) - Y_{n-1}) + E(Y_2|\mathscr{F}_1),$$

$$\forall m \geqslant 2,$$

是 $\mathscr{F}_{m-1}$ 可测的。

由推论 2.7.4 和式(2.7.9)可知 Z_m 在

$$\{\sup_{n\geqslant 1} Y_n < \infty, \quad \sup_{n\geqslant 1} U_n < \infty\}$$

上 a.s. 收敛。

因为 $Z_n + S_n$ 是下鞅，且 $S_n \leqslant Y_n$

$$Z_n + S_n \leqslant Z_n + Y_n \leqslant U_n$$

因此 $Z_n + S_n$ 在 $\{\sup U_n < \infty\}$ 上 a.s. 收敛。故 S_n 在集(2.7.8)上 a.s. 收敛。 ∎

推论 2.7.6 设 $(S_n, \mathscr{F}_n)$ 为下鞅，$p \geqslant 1$

(1) 若 $E(S_n^+)^p < \infty$，则 S_n 在

$$\left\{\sum_{2}^{\infty} (E((S_n^+)^p|\mathscr{F}_{n-1}) - (S_{n-1}^+)^p) < \infty\right\} \quad (2.7.10)$$

上 a.s. 收敛。

(2) 若 $E|S_n|^p < \infty$，则 S_n 在

$$\left\{\sum_{2}^{\infty} (E(|S_n|^p|\mathscr{F}_{n-1}) - |S_{n-1}|^p) < \infty\right\} \quad (2.7.11)$$

上 a.s. 收敛。

证明 因为

$$S_n \leqslant S_n^+ \leqslant (S_n^+)^p + 1, \quad S_n \leqslant |S_n| \leqslant |S_n|^p + 1$$

由推论 2.7.5，本推论得证。 ∎

定理 2.7.6 设 $(S_n, \mathscr{F}_n)$ 为鞅，$E|S_n| < \infty$，$X_1 = S_1$，$X_n = S_n - S_{n-1}$，$\forall n \geqslant 2$，以及

$$A = \left\{\omega: \sum_{n=2}^{\infty} E(|X_n|^2 \chi_{[|X_n| \leqslant a_n]}\right.$$

$$+ |X_n|\chi_{[|X_n|>a_n]}|\mathscr{F}_{n-1}) < \infty\} \quad (2.7.12)$$

其中 $a_n \geqslant c > 0$，c 为常数，那么 S_n 在 A 上 a.s. 收敛

证明 令 $X'_n = X_n\chi_{[|X_n|\leqslant a_n]}$，于是

$$|E(X'_n|\mathscr{F}_{n-1})| = |E(X_n\chi_{[|X_n|>a_n]}|\mathscr{F}_{n-1})|$$
$$\leqslant E(|X_n|\chi_{[|X_n|>a_n]}|\mathscr{F}_{n-1})$$

那么在 A 上，有

$$\sum_1^\infty E(X'_n|\mathscr{F}_{n-1}) < \infty, \text{ a.s.} \quad (2.7.13)$$

因为

$$\sum_1^\infty P(X'_n \neq X_n|\mathscr{F}_{n-1}) = \sum_1^\infty P(|X_n| > a_n|\mathscr{F}_{n-1})$$

$$\leqslant c^{-1}\sum_1^\infty E(|X_n|\chi_{[|X_n|>a_n]}|\mathscr{F}_{n-1}) < \infty, \forall \omega \in A$$

所以

$$\sum_1^\infty P(X'_n \neq X_n) < \infty$$

在 A 上成立，即

$$P(A\cap[X'_n \neq X_n, \text{i.o.}]) = 0 \quad (2.7.14)$$

再令

$$Y_n = X'_1 + \cdots + X'_n - E(X'_2|\mathscr{F}_1)$$
$$- \cdots - E(X'_n|\mathscr{F}_{n-1})$$

于是

$$E(Y_n|\mathscr{F}_{n-1}) = E\left\{\left(\sum_{i=1}^n X'_i - \sum_{i=2}^n E(X'_i|\mathscr{F}_{i-1})\right)\middle|\mathscr{F}_{n-1}\right\}$$

$$= \sum_{i=1}^{n-1} X'_i - \sum_{i=2}^n E(X'_i|\mathscr{F}_{i-1}) + E(X'_n|\mathscr{F}_{n-1}) = Y_{n-1}$$

故 $(Y_n, \mathscr{F}_n)$ 是鞅，且

$$E(Y_n^2|\mathscr{F}_{n-1})-Y_{n-1}^2$$
$$=E\{[Y_{n-1}+X_n'-E(X_n'|\mathscr{F}_{n-1})]^2|\mathscr{F}_{n-1}\}-Y_{n-1}^2$$
$$=E[(X_n'-E(X_n'|\mathscr{F}_{n-1}))^2|\mathscr{F}_{n-1}]$$
$$=E((X_n')^2|\mathscr{F}_{n-1})-[E(X_n'|\mathscr{F}_{n-1})]^2$$

而且

$$E(Y_n^2|\mathscr{F}_{n-1})-Y_{n-1}^2\leqslant E(X_n'^2|\mathscr{F}_{n-1})$$
$$\leqslant E(X_n^2\chi_{[|X_n|\leqslant c]}|\mathscr{F}_{n-1})$$

因此在A上有

$$\sum_2^\infty(E(Y_n^2|\mathscr{F}_{n-1})-Y_{n-1}^2)<\infty,\ \text{a.s.}$$

由推论 2.7.6(2) 可知，Y_n 在A上 a.s. 收敛。再由式 (2.7.13) 和 (2.7.14) 可得 S_n 在A上 a.s. 收敛。 ▌

推论 2.7.7 设 $1\leqslant p\leqslant 2$，记

$$A=\left\{\sum_{i=2}^\infty E(|X_i|^p|\mathscr{F}_{i-1})<\infty\right\} \tag{2.7.15}$$

则 S_n 在A上 a.s. 收敛。

证明 因为

$$\sum_{i=2}^\infty E(|X_i|^2\chi_{[|X_i|\leqslant 1]}+|X_i|\chi_{[|X_i|>1]}|\mathscr{F}_{i-1})$$
$$\leqslant\sum_{i=2}^\infty E(|X_i|^p|\mathscr{F}_{i-1})<\infty$$

由定理 2.7.6，本推论得证。 ▌

习　题　2.7

1. 试写出下鞅的上穿不等式。
2. 试写出与定理 2.7.1 相应的下鞅收敛定理。
3. 设 $(X_n,\mathscr{F}_n)$ 为鞅，$EX_n^2<\infty$，$\forall n\geqslant 0$。证明如果存在 $n_0<\infty$

使得 $X_{n_0+k}=X_{n_0}, \forall k\geq 0$，那么

$$E|\sup_n X_n|^2\leq q^2E[X_{n_0}^2]$$

其中 q 为某常数。

4. 证明引理 2.7.1。

2.8 几乎上鞅和新息序列的收敛定理

提要 本节证明了一种几乎上鞅列的收敛定理、Toeplitz 引理、Kronecker 引理和新息序列的极限定理。

设 $(\Omega,\mathscr{F},P)$ 为概率空间，$\mathscr{F}_1\subset\mathscr{F}_2\subset\cdots\subset\mathscr{F}$ 是 $\mathscr{F}$ 的完备子 σ 代数列。对于 $n=1,2,\cdots$，Z_n,η_n,ξ_n,ζ_n 都是非负的 $\mathscr{F}_n$ 可测随机变量序列。

定理 2.8.1 若

$$E(Z_{n+1}|\mathscr{F}_n)\leqslant Z_n(1+\eta_n)+\xi_n-\zeta_n. \quad (2.8.1)$$

则在

$$\left\{\omega\left|\sum_1^\infty\eta_n<\infty,\ \sum_1^\infty\xi_n<\infty\right.\right\}$$

集上 $\lim\limits_{n\to\infty}Z_n$ 存在且有限，而且 $\sum\limits_1^\infty\zeta_n<\infty$，a.s.。

证明 令

$$A=\left\{\omega\left|\sum_1^\infty\eta_n<\infty,\ \sum_1^\infty\xi_n<\infty\right.\right\}$$

定义

$$Z_n'=Z_n\prod_{k=1}^{n-1}(1+\eta_k)^{-1},\quad \xi_n'=\xi_n\prod_{k=1}^{n}(1+\eta_k)^{-1}$$

和

$$\zeta_n' = \zeta_n \prod_{k=1}^{n} (1 + \eta_k)^{-1}$$

因为 $\eta_k \geqslant 0$, $\forall k \geqslant 1$, 所以上面定义的三个等式均有意义.

由式 (2.8.1) 可得

$$E(Z_{n+1}' | \mathscr{F}_n) \leqslant Z_n' + \xi_n' - \zeta_n' \tag{2.8.2}$$

在 A 上

$$\sum \xi_n' \leqslant \sum \xi_n < \infty \tag{2.8.3}$$

令

$$U_n = Z_n' - \sum_{k=1}^{n-1} (\xi_k' - \zeta_k') \tag{2.8.4}$$

对 $a > 0$

$$t = \inf\left\{n: \sum_1^n \xi_k' > a\right\} \tag{2.8.5}$$

在集 $\{t > n\}$ 上,由式 (2.8.2) 和 (2.8.4) 得

$$E(U_{n+1} | \mathscr{F}_n) = E\left(Z_{n+1}' - \sum_1^n (\xi_k' - \zeta_k') | \mathscr{F}_n \right)$$

$$\leqslant Z_n' + \xi_n' - \zeta_n' - \sum_1^n \xi_k' + \sum_1^n \zeta_k' = U_n$$

而且有

$$\begin{aligned} E(U_{t\wedge(n+1)} | \mathscr{F}_n) &= U_t \chi_{[t \leqslant n]} + E(U_{n+1} | \mathscr{F}_n) \chi_{[t > n]} \\ &\leqslant U_t \chi_{[t \leqslant n]} + U_n \chi_{[t > n]} = U_{t\wedge n} \end{aligned}$$

所以 $\{U_{t\wedge n}, \mathscr{F}_n\}$ 是上鞅. 又因为

$$U_{t\wedge n} \geqslant - \sum_{k=1}^{t\wedge n-1} \xi_k' \geqslant -a$$

因此

$$EU_{t\wedge n}^- \leqslant a$$

由定理 2.7.1 知 $\lim\limits_{n\to\infty} U_{t\wedge n}$ 存在且有限,即在

$$\{\omega \mid t=\infty\}=\left\{\omega \,\middle|\, \sum_{1}^{\infty} \xi_n' \leqslant a\right\}$$

上 $\lim\limits_{n\to\infty} U_n$ 存在且有限。

因为 a 是任意的，因此在

$$\left\{\omega \,\middle|\, \sum_{1}^{\infty} \xi_n' < \infty\right\}$$

上 $\lim\limits_{n} U_n$ 存在且有限。

由于式(2.8.4)在 $\{\sum \xi_n' < \infty\}$ 上 $\lim Z_n'$ 存在、有限且 $\sum \zeta_n' < \infty$. 因此，由式(2.8.3)和

$$Z_n = Z_n' \prod_{1}^{n-1} (1+\eta_k)$$

以及不等式

$$\zeta_n \leqslant \zeta_n' \prod_{1}^{\infty} (1+\eta_k)$$

可知，在 A 上 $\lim\limits_{n} Z_n$ 存在、有限且 $\sum\limits_{1}^{\infty} \zeta_n < \infty$ a.s. ∎

推论 2.8.1 设对于 $n=1, 2, \cdots, Z_n, \xi_n, \zeta_n$ 是关于 $\mathscr{F}_n$ 可测的随机变量序列，$\mathscr{F}_n$ 是 $\mathscr{F}$ 的增子 σ 代数，假定

$$E(Z_{n+1} \mid \mathscr{F}_n) \leqslant Z_n + \xi_n - \zeta_n \tag{2.8.6}$$

那么，当 $\sum\limits_{1}^{\infty} \xi_n < \infty$ 时有

(1) $\lim\limits_{n} Z_n = Z$, a.s. Z 有限;

(2) $\sum\limits_{1}^{\infty} \zeta_n < \infty$, a.s..

证明 当 $\eta_n = 0$ 时得推论。 ∎

注意：(1) 这里的条件 $\sum\xi_n<\infty$ a.s. 比文献[13]中定理 2.7.4 条件放宽了，但结论是相同的。

(2) 定理 2.8.1 的证明引自文献[15]。

(3) $(Z_n, n\in N)$ 满足式(2.8.1)，Z_n 称为几乎上鞅，故定理 2.8.1 和推论 2.8.1 称为几乎上鞅收敛定理。

引理 2.8.1 (Toeplitz 引理) 设 $\Phi_{n,k}$, $k=1,2,\cdots$ 为 $p\times p$ 实数矩阵列，对某一固定的 k 有

$$\lim_{n\to\infty}\Phi_{n,k}=0\ (\text{零矩阵})$$

对所有的 n 有

$$\sum_{k=1}^{\infty}\|\Phi_{n,k}\|\leqslant c_1(\text{常数})<\infty$$

设

$$y_n=\sum_{k=1}^{\infty}\Phi_{n,k}x_k$$

其中 $\{x_k\}$ 是 p 维实数向量序列，那么

(1) 由 $\lim\limits_{k\to\infty}x_k=0$ 可推出 $\lim\limits_{n\to\infty}y_n=0$；

(2) 若 $\lim\limits_{n\to\infty}\sum\limits_{k=1}^{\infty}\Phi_{n,k}=I$，由 $\lim\limits_{k\to\infty}x_k=x<\infty$ 可推出

$$\lim_{n\to\infty}y_n=x$$

证明 (1) 由已知条件 $\lim\limits_{k\to\infty}x_k=0$ 知，对任何 $\varepsilon>0$，存在 $k(\varepsilon)$ 使得对于所有 $k>k(\varepsilon)$，有

$$\|x_k\|<\varepsilon/c_1$$

由于

$$y_n=\sum_{k=1}^{\infty}\Phi_{n,k}x_k=\sum_{k=1}^{k(\varepsilon)}\Phi_{n,k}x_k+\sum_{k=k(\varepsilon)+1}^{\infty}\Phi_{n,k}x_k$$

利用范数不等式可得

$$\|y_n\| \leqslant \left\|\sum_{k=1}^{k(\varepsilon)} \Phi_{nk}x_k\right\| + \sum_{k=k(\varepsilon)+1}^{\infty} \|\Phi_{n,k}\|\|x_k\|$$

$$\leqslant \left\|\sum_{k=1}^{k(\varepsilon)} \Phi_{nk}x_k\right\| + \varepsilon$$

再利用 $\Phi_{n,k} \xrightarrow[n\to\infty]{} 0$ 得

$$\lim_{n\to\infty} \|y_n\| \leqslant \varepsilon$$

由 ε 的任意性可得

$$\lim_{n\to\infty} y_n = 0$$

(2) 因为

$$y_n = \sum_{k=1}^{\infty} \Phi_{nk}x + \sum_{k=1}^{\infty} \Phi_{nk}(x_k - x)$$

利用

$$\lim_{n\to\infty} \sum_{k=1}^{\infty} \Phi_{nk} = I \text{ 和 } \lim_{k\to\infty} (x_k - x) = 0$$

以及本引理(1)可得

$$\lim_{n\to\infty} y_n = x + 0 = x$$ ▮

注意：在以上引理中 $\|\cdot\|$ 表示矩阵的范数，如 $\|A\| = \{\lambda_{max}A^TA\}^{1/2}$，$\lambda_{max}(\cdot)$ 表示矩阵的最大特征值。

引理2.8.2 (Kronecker 引理) 如果

$$S_n = \sum_{k=1}^{n} A_k,\quad A_k \text{ 是 } p \times p \text{ 半正定阵} \tag{2.8.7}$$

$$v_n = \sum_{k=1}^{n} S_k^{-1}x_k,\quad \forall \text{某一个 } p \text{ 维向量 } x_k \tag{2.8.8}$$

那么由

$$\lim_{n\to\infty} v_n = v, \quad \|v\| < \infty \tag{2.8.9}$$

$$\lim_{n\to\infty} S_n^{-1} A_k = 0, \forall k > 0 \tag{2.8.10}$$

$$\sum_{k=1}^{n} \|S_n^{-1} A_k\| < \infty, \forall n > 0 \tag{2.8.11}$$

可导出

$$\lim_{n\to\infty} S_n^{-1} \sum_{k=1}^{n} x_k = 0 \tag{2.8.12}$$

证明 本引理主要利用 Toeplitz 引理来证明。

令 $v_0 = 0$，由式(2.8.8)得

$$v_k - v_{k-1} = S_k^{-1} x_k$$

即

$$S_k(v_k - v_{k-1}) = x_k, \quad \forall k > 0$$

因此

$$S_n^{-1} \sum_{k=1}^{n} x_k = S_n^{-1} \sum_{k=1}^{n} S_k(v_k - v_{k-1})$$

$$= v_n - S_n^{-1} S_n v_n + S_n^{-1} \sum_{k=1}^{n} S_k(v_k - v_{k-1})$$

$$= v_n - S_n^{-1} \left(S_n v_n - \sum_{k=1}^{n} S_k(v_k - v_{k-1}) \right)$$

$$= v_n - S_n^{-1} \sum_{k=1}^{n} (S_k - S_{k-1}) v_{k-1}$$

$$= v_n - S_n^{-1} \sum_{k=1}^{n} A_k v_{k-1}$$

定义

$$\Phi_{nk} = \begin{cases} S_n^{-1} A_k, & \forall 1 < k \leqslant n \\ 0, & \forall k > n \end{cases}$$

所以

$$\lim_{n\to\infty}\Phi_{nk}=0$$

由条件（2.8.11）得

$$\sum_{k=1}^{n}\|\Phi_{nk}\|=\sum_{k=1}^{n}\|S_n^{-1}A_k\|<\infty,\ \forall n\geq 1$$

应用 Toeplitz 引理，有

$$\lim_{n\to\infty}S_n^{-1}\sum_{k=1}^{n}A_k v_{k-1}=\lim_{n\to\infty}\sum_{k=1}^{\infty}\Phi_{nk}v_{k-1}=v$$

于是

$$\lim_{n\to\infty}S_n^{-1}\sum_{k=1}^{n}x_k=\lim_{n\to\infty}v_n-\lim_{n\to\infty}S_n^{-1}\sum_{k=1}^{n}A_k v_{k-1}$$

$$=v-v=0$$

定义 2.8.1 设 $\{\mathscr{F}_n\}$ 是一递增 σ 代数序列，即

$$\mathscr{F}_0\subset\mathscr{F}_1\subset\cdots\subset\mathscr{F}$$

$\{\varepsilon_n\}$ 适应于 $\{\mathscr{F}_n\}$. 如果

$$E(\varepsilon_{n+1}|\mathscr{F}_n)=0,\forall n>0$$

则称 $\{\varepsilon_n\}$ 为鞅差序列或新息序列.

定理 2.8.2 设 $\{\varepsilon_n\}$ 是一个适应于 $\{\mathscr{F}_n\}$ 的新息序列，$\{X_n\}$ 定义为

$$X_n=\sum_{t=1}^{n}S_t^{-1}\varepsilon_t,\ \ n=1,2,\cdots$$

$$X_0=0$$

其中

$$S_t=\sum_{k=1}^{t}\phi_k$$

ϕ_k 是半正定对称方阵. S_t 关于 $\mathscr{F}_{t-1}$ 可测，且

$$\left\| E\left(\sum_{i=1}^{n} S_i^{-1}\varepsilon_i\varepsilon_i^T S_i^{-1}\right)\right\| < c\ (\text{常数}) < \infty$$

$$\lim_{n\to\infty} S_n^{-1}\psi_k = 0,\ \forall k \geqslant 1\ \text{a.s.}$$

$$\sum_{k=1}^{n} \|S_i^{-1}\psi_k\| < \infty, \forall n \geqslant 1\ \text{a.s.}$$

那么

(1) $(X_n, \mathscr{F}_n)$ 是鞅；

(2) X_n 几乎必然收敛；

(3) $S_n^{-1}\sum_{i=1}^{n}\varepsilon_i \xrightarrow[n\to\infty]{} 0$, a.s..

证明 (1) $E(X_n|\mathscr{F}_{n-1}) = E(S_n^{-1}\varepsilon_n + X_{n-1}|\mathscr{F}_{n-1})$

$$= E(S_n^{-1}\varepsilon_n|\mathscr{F}_{n-1}) + X_{n-1}$$

由于假设 S_n 关于 $\mathscr{F}_{n-1}$ 可测，由引理 2.4.4，有

$$E(S_n^{-1}\varepsilon_n|\mathscr{F}_{n-1}) = S_n^{-1}E(\varepsilon_n|\mathscr{F}_{n-1}) = 0$$

因此 $(X_n, \mathscr{F}_n)$ 是鞅。

(2) 考察 $E(X_nX_n^T)$ 是否有界：

$$E(X_nX_n^T) = E\left(\sum_{k=1}^{n} S_k^{-1}\varepsilon_k\right)\left(\sum_{j=1}^{n} S_j^{-1}\varepsilon_j\right)^T$$

$$= E\left(\sum_{i=1}^{n} S_i^{-1}\varepsilon_i\varepsilon_i^T S_i^{-1} + \sum_{k=1}^{n}\sum_{j=k+1}^{n} S_k^{-1}\varepsilon_k\varepsilon_j^T S_j^{-1}\right.$$

$$\left. + \sum_{k=1}^{n}\sum_{j=1}^{k-1} S_k^{-1}\varepsilon_k\varepsilon_j^T S_j^{-1}\right) = E\left\{\sum_{i=1}^{n} S_i^{-1}\varepsilon_i\varepsilon_i^T S_i^{-1}\right.$$

$$+ \sum_{k=1}^{n}\sum_{j=k+1}^{n} E(S_k^{-1}\varepsilon_k\varepsilon_j^T S_j^{-1}|\mathscr{F}_{j-1})$$

$$\left. + \sum_{k=1}^{n}\sum_{j=1}^{k-1} E(S_k^{-1}\varepsilon_k\varepsilon_j^{-1} S_j^{-1}|\mathscr{F}_{k-1})\right\}$$

$$= E\left[\sum_{t=1}^{n} S_t^{-1}\varepsilon_t\varepsilon_t^T S_t^{-1}\right]$$

因为后两项都是零.

根据假定

$$\|E(X_n X_n^T)\| < c \text{（常数）} < \infty$$

由推论 2.7.2 有

$$\lim_{n\to\infty} X_n = X, \text{ a.s.}, \text{ 且 } X \text{ 可积}$$

(3) 由于假设

$$\lim_{n\to\infty} S_n^{-1}\psi_k = 0, \ \forall k \geqslant 1, \text{ a.s.}$$

$$\sum_{k=1}^{n} \|S_n^{-1}\psi_k\| < \infty, \ \forall n \geqslant 1, \text{ a.s.}$$

根据(2)

$$\lim_{n\to\infty} X_n = X, \text{ a.s.}$$

因此 Kronecker 引理的所有条件都满足,由此可得

$$\lim_{n\to\infty} S_n^{-1}\sum_{t=1}^{n} \varepsilon_t = 0, \text{ a.s.}$$

习 题 2.8

1. 设 $(X_n^{(i)}, n\in N)$ $(i=1, 2)$ 是两个正上鞅, ν 为停时且使 $X_\nu^{(1)} \geqslant X_\nu^{(2)}$, $\omega\in\{\nu<\infty\}$, 证明

$$X_n(\omega) = \begin{cases} X_n^{(1)}(\omega), & \text{当 } n<\nu(\omega) \\ X_n^{(2)}(\omega), & \text{当 } n\geqslant\nu(\omega) \end{cases}, n\in N$$

确定了一个正上鞅.

2. 证明每一个正上鞅必定几乎必然收敛且

$$X_\infty = \lim_n X_n, \text{a.s.}$$

满足

$$E(X_\infty|\mathscr{F}_n)\leqslant X_n, n\in N$$

3. 若 $(X_n, n\in N)$ 是正上鞅，ν 是 $(\Omega,\mathscr{F},\mathscr{F}_n,P)$ 上的停时，令

$$X'_n = E(X_{\nu\wedge n}|\mathscr{F}_n), n\in N$$

证明 $(X'_n, n\in N)$ 是正上鞅.

2.9 上鞅列的分解

提要 本节讨论离散时间鞅的 Doob 分解和 Riesz 分解，并对上鞅的结构作深入的研究。

随机序列 $(X_n, \mathscr{F}_n)$ 称为递增随机序列，若

$$0 = X_1 \leqslant X_2 \leqslant \cdots$$

则存在 $X_\infty = \lim\limits_n X_n$, a.s.. 若 X_∞ 可积，则递增序列是可积的。若 X_∞ 可积，即 $\{X_n\}$ 为可积递增随机序列，那么

(1) $EX_n \leqslant EX_\infty < \infty$，即对一切 n，一致有界；

(2) 因为 $EX_\infty < \infty$，对任给 $\varepsilon > 0$，存在 $\delta > 0$，只要 $A\in\mathscr{F}$，$p(A)\leqslant\delta$ 就有

$$\int_A X_n dp \leqslant \int_A X_\infty dp < \varepsilon, \quad \forall n \geqslant 1$$

根据 $\{X_n\}$ 一致可积的充分必要条件以及定理 2.4.3 可知 $\{X_n\}$ 一致可积.

因此，可积递增随机序列是一致可积的.

定义 2.9.1 设 $(X_n, \mathscr{F}_n)$ 是随机序列，若对于每一个 $n\geqslant 1$，X_{n+1} 是 $\mathscr{F}_n$ 可测的，则称 $(X_n,\mathscr{F}_n)$ 是可料的.

定义 2.9.2 上鞅 $(X_n, \mathscr{F}_n)$ 如能表示成

$$X_n = m_n - Z_n, \quad \forall n \geqslant 1 \tag{2.9.1}$$

其中 $(m_n, \mathscr{F}_n)$ 为鞅，$(Z_n, \mathscr{F}_n)$ 为递增随机序列，则称

式(2.9.1)为上鞅的 Doob 分解。

定理 2.9.1 任一上鞅具有 Doob 分解。如果分解中的递增序列 Z_n 可料，那么分解唯一。

证明 定义

$$m_1 = X_1$$

$$m_{n+1} - m_n = X_{n+1} - E(X_{n+1}|\mathscr{F}_n), \quad \forall n \geqslant 1$$

并且令

$$Z_1 = 0, \quad Z_{n+1} - Z_n = X_n - E(X_{n+1}|\mathscr{F}_n) \ \forall n \geqslant 1$$

由于

$$E(m_{n+1} - m_n|\mathscr{F}_n) = E(X_{n+1}|\mathscr{F}_n) - E(X_{n+1}|\mathscr{F}_n) = 0$$

因此

$$E(m_{n+1}|\mathscr{F}_n) = m_n$$

即 $(m_n, n \in \boldsymbol{N})$ 是鞅。

又因为 $(X_n, n \in \boldsymbol{N})$ 是上鞅，所以

$$Z_{n+1} - Z_n \geqslant 0, \quad \forall n \geqslant 1$$

$(Z_n, n \in \boldsymbol{N})$ 是随机递增序列。

容易验证

$$X_2 = m_2 - Z_2$$

根据数学归纳法。假设

$$X_{n-1} = m_{n-1} - Z_{n-1} \tag{2.9.2}$$

由定义

$$X_n = m_n - m_{n-1} + E(X_n|\mathscr{F}_{n-1}) \tag{2.9.3}$$

把式(2.9.2)代入式(2.9.3)，得

$$\begin{aligned} X_n &= m_n - (X_{n-1} + Z_{n-1}) + E(X_n|\mathscr{F}_{n-1}) \\ &= m_n - (X_{n-1} - E(X_n|\mathscr{F}_{n-1})) - Z_{n-1} \\ &= m_n - Z_n + Z_{n-1} - Z_{n-1} \\ &= m_n - Z_n \end{aligned}$$

因此式(2.9.1)成立。

证明分解的唯一性：

设有另一个鞅（m'_n，$\mathscr{F}_n$）以及可料增随机序列（Z'_n，$\mathscr{F}_n$）使得

$$X_n = m'_n - Z'_n, \ \forall n \geqslant 1$$

那么

$$Z'_{n+1} - Z'_n = (m'_{n+1} - m'_n) + (X_n - X_{n+1})$$

对$\mathscr{F}_n$取条件数学期望：

$$Z'_{n+1} - Z'_n = X_n - E(X_{n+1}|\mathscr{F}_n) = Z_{n+1} - Z_n$$

因为 $Z_1 = Z'_1 = 0$，所以 $Z'_n = Z_n$，$\forall n \geqslant 1$。从而有 $m'_n = m_n$，$\forall n \geqslant 1$ 成立。 ∎

Doob 分解表明上鞅等于鞅与可料增序列之差，利用这一研究可以由上鞅的收敛性质推出鞅的收敛性质，或利用鞅的收敛性质推出上鞅的收敛性质，因为递增序列的收敛性比较容易判别。

定理 2.9.2 设（X_n，$n \in \boldsymbol{N}$）为一致可积上鞅，则其 Doob 分解中的鞅和递增序列也一致可积。

证明 设（X_n，$n \in \boldsymbol{N}$）的 Doob 分解为

$$X_n = m_n - Z_n$$

由于 $\{X_n\}$ 一致可积，故

$$\sup_n E|X_n| < \infty \ (\text{一致有界})$$

设

$$\lim_n Z_n = Z_\infty$$

于是有

$$EZ_\infty = \lim_n EZ_n = Em_n - \lim_{n\to\infty} EX_n < \infty$$

因此 $\{Z_n\}$ 为可积递增随机序列，而且 $\{Z_n\}$ 一致可积。

由于 $m_n = X_n + Z_n$，于是鞅 $\{m_n\}$ 也一致可积。

定义 2.9.3 设 $(X_n, \mathscr{F}_n)$ 是非负上鞅，且

$$\lim_n EX_n = 0$$

即 $X_n \xrightarrow{L_1} 0$，此 $(X_n, \mathscr{F}_n)$ 称为位势。

推论 2.9.1 位势是一致可积上鞅。

证明留作习题。

定理2.9.3 随机序列 $(X_n, \mathscr{F}_n)$ 是位势的充分必要条件为：存在可积递增随机序列 $(Z_n, \mathscr{F}_n)$ 使得

$$X_n = E(Z_\infty | \mathscr{F}_n) - Z_n, \quad \forall n \geqslant 1 \tag{2.9.4}$$

若 $(Z_n, n \in \boldsymbol{N})$ 可料，则式(2.9.4)表示唯一。

证明 充分性：若式(2.9.4)成立，那么对任意的 $n \geqslant 1$，有

$$\begin{aligned} E(X_{n+1}|\mathscr{F}_n) &= E(Z_\infty|\mathscr{F}_n) - E(Z_{n+1}|\mathscr{F}_n) \\ &\leqslant E(Z_\infty|\mathscr{F}_n) - E(Z_n|\mathscr{F}_n) \\ &\doteq E(Z_\infty|\mathscr{F}_n) - Z_n = X_n \end{aligned}$$

且

$$X_n = E(Z_\infty - Z_n|\mathscr{F}_n) \geqslant 0$$

可以证明上鞅 $X_n \xrightarrow{L_1} 0$（留作习题），故 $(X_n, \mathscr{F}_n)$ 是位势。

必要性：设 $(X_n, n \in \boldsymbol{N})$ 是位势，它的 Doob 分解为

$$X_n = m_n - Z_n, \quad \forall n \geqslant 1$$

由定理 2.9.2 知，位势是一致可积上鞅，因此其 Doob 分解中的鞅也一致可积。由鞅定义得到

$$\begin{aligned} m_n &= E(m_{n+1}|\mathscr{F}_n) = E(E(m_{n+2}|\mathscr{F}_{n+1})|\mathscr{F}_n) \\ &= E(m_{n+2}|\mathscr{F}_n) = \cdots = E(m_\infty|\mathscr{F}_n) \end{aligned}$$

且

$$m_\infty = \lim_n m_n, \quad \text{a.s.}$$

因此 $(Z_n, n \in \boldsymbol{N})$ 是可积递增随机序列，X_n 是位势；

$$\lim_n X_n = m_\infty - Z_\infty = 0, \ \text{a.s.}$$

而且

$$m_n = E(Z_\infty | \mathscr{F}_n), \ \forall n \geqslant 1$$

即式(2.9.4)成立,唯一性可由定理2.9.1得到。 ∎

定义 2.9.4 上鞅 $(X_n, n \in \boldsymbol{N})$ 可表示为

$$X_n = Y_n + Z_n, \ \forall n \geqslant 1 \tag{2.9.5}$$

其中 $(Y_n, n \in \boldsymbol{N})$ 是鞅,$(Z_n, n \in \boldsymbol{N})$ 是位势,称 $(X_n, n \in \boldsymbol{N})$ 具有 Riesz 分解。

引理 2.9.1 设 $(X_n, \mathscr{F}_n)$ 是上鞅,如果它具有 Riesz 分解,那么其 Riesz 分解必唯一。

证明 设 $X_n = Y_n + Z_n$, $X_n = Y'_n + Z'_n$ 是上鞅 $(X_n, \mathscr{F}_n)$ 的两个 Riesz 分解。因为 Y_n, Y'_n 均为鞅,因此 $(Y_n - Y'_n, \mathscr{F}_n)$ 是鞅.

又

$$Y_n - Y'_n = Z'_n - Z_n, \ \forall n \geqslant 1$$

以及 Z'_n 和 Z_n 均为位势,因此由推论2.9.1,$\{Z'_n - Z_n\}$ 一致可积,$\{Y_n - Y'_n\}$ 一致可积且

$$Y_n - Y'_n \xrightarrow{L_1, \text{a.s}} 0$$

由本章定理2.7.2和推论2.7.1可知

$$E(Y_\infty - Y'_\infty | \mathscr{F}_n) = Y_n - Y'_n = 0$$

即对一切 n 有 $Y_n = Y'_n$, a.s. 从而有 $Z_n = Z'_n$, a.s.. ∎

定理 2.9.4 上鞅 $(X_n, \mathscr{F}_n)$ 具有 Riesz 分解的充分必要条件为

$$\lim_n EX_n > -\infty \tag{2.9.6}$$

证明 必要性:设 $(X_n, \mathscr{F}_n)$ 具有 Riesz 分解

$$X_n = Y_n + Z_n$$

那么对充分大的 n 有

$$EX_n \geqslant EY_n = EY_1$$

因此式(2.9.6)成立.

充分性：如果式(2.9.6)成立，因为$(X_n, n\in \boldsymbol{N})$是上鞅，它的 Doob 分解为

$$X_n = m_n - a_n$$

于是

$$Ea_\infty = \lim_n Ea_n = Em_n - \lim_n EX_n < \infty$$

因此$(a_n, n\in \boldsymbol{N})$是可积递增随机序列.

令

$$Z_n = E(a_\infty|\mathscr{F}_n) - a_n, \quad \forall n\geqslant 1 \tag{2.9.7}$$

$$Y_n = m_n - E(a_\infty|\mathscr{F}_n), \quad \forall n\geqslant 1 \tag{2.9.8}$$

那么$(Z_n, n\in \boldsymbol{N})$为位势，$(Y_n, n\in \boldsymbol{N})$为鞅且

$$X_n = Y_n + Z_n, \quad \forall n\geqslant 1$$

即$(X_n, n\in \boldsymbol{N})$具有 Riesz 分解. ∎

推论 2.9.2 设$(X_n, n\in \boldsymbol{N})$为非负上鞅，那么$(X_n, n\in \boldsymbol{N})$具有 Riesz 分解，且分解中的鞅$(Y_n, n\in \boldsymbol{N})$也非负.

证明留作习题.

推论 2.9.3 设$(X_n, n\in \boldsymbol{N})$是一致可积上鞅，那么$(X_n, n\in \boldsymbol{N})$具有下列 Riesz 分解:

$$X_n = E(X_\infty|\mathscr{F}_n) + (X_n - E(X_\infty|\mathscr{F}_n)), \quad \forall n\geqslant 1$$

其中

$$\lim_n X_n = X_\infty$$

证明留作习题.

习　题　2.9

1. 证明推论 2.9.1.

2. 证明式(2.9.7)和(2.9.8)成立.

3.证明推论2.9.2.

4.证明推论2.9.3.

5.证明由式(2.9.4)定义的 $X_n \xrightarrow{L_1} 0$.

2.10 鞅差中心极限定理

提要 本节讨论双下标随机序列的中心极限定理。首先，我们给出一个双下标相依随机序列的中心极限定理，然后证明鞅差中心极限定理。

相依随机序列的中心极限定理对研究参数辨识、传递函数辨识的渐近性质，是很重要的数学工具。本节主要介绍 D. L. McLeish 的中心极限定理。

设 $(\Omega, \mathscr{F}, P)$ 是概率空间. $\{X_{n,j};\ j=1,\cdots,k_n;\ n=1,2,\cdots\}$ 是定义在 $(\Omega, \mathscr{F}, P)$ 上的双下标随机序列。令

$$S_n = \sum_{j=1}^{k_n} X_{n,j} \tag{2.10.1}$$

设 $\{\mathscr{F}_{n,j}: 0 \leqslant j \leqslant k_n\}$ 是 $\mathscr{F}$ 的子 σ 代数的双下标序列，对每一个 n 以及 $1 \leqslant j \leqslant k_n$, $X_{n,j}$ 是 $\mathscr{F}_{n,j}$ 可测的且 $\mathscr{F}_{n,j-1} \subset \mathscr{F}_{n,j}$. 设 i 是虚单位，即 $i^2=-1$, t 为实数，对于每一个正整数 n, 用下式定义随机变量.

$$T_n = \prod_{j=1}^{k_n} (1 + itX_{n,j}) \tag{2.10.2}$$

下面给出一般相依随机序列的中心极限定理。

定理 2.10.1 (McLeish) 假设对于一切实数 t 有

(1) $ET_n \xrightarrow[n\to\infty]{} 1$;

(2) $\{T_n\}$ 一致可积;

(3) $\sum_{j=1}^{k_n} X_{n,j}^2 \xrightarrow[n\to\infty]{P} 1$;

(4) $\max_{j\leqslant k_n}|X_{n,j}| \xrightarrow[n\to\infty]{P} 0$。

那么式 (2.10.1) 的 S_n 分布收敛于 $N(0,1)$ 正态分布。

证明 设 S_n 的特征函数为 EI_n。欲证明 S_n 分布收敛于 $N(0,1)$，只要证明 EI_n 收敛于 $e^{-\frac{t^2}{2}}$ 即可。因为 $e^{-\frac{t^2}{2}}$ 是 $N(0,1)$正态随机变量的特征函数。根据特征函数的极限定理：设特征函数列 $\{f_n(t)\}$ 收敛于某一连续函数 $f(t)$，则相应的分布函数 $\{F_n(x)\}$ 弱收敛于某一分布函数 $F(x)$，即在 $F(x)$的每一个连续点上收敛于 $F(x)$，且

$$\lim_{n\to\infty} F_n(\pm\infty) = F(\pm\infty)$$

并且有

$$f(t) = \int e^{itx} dF(x)$$

先看下列事实：因为

$$\log(1+z) = z - \frac{1}{2}z^2 + o(z^2)$$

即

$$1 + z = e^{z-\frac{1}{2}z^2+o(z^2)}$$

令

$$z = ix$$

于是有

$$(1+ix) = e^{ix+\frac{1}{2}x^2+\gamma(x)}$$

其中 $\gamma(x) = O(x^3)$，所以

$$e^{ix} = (1+ix)e^{-\frac{1}{2}x^2+\gamma(x)}$$

再令

$$I_n \triangleq e^{itS_n} = e^{it\sum_{j=1}^{k_n} X_{n,j}} = \prod_{j=1}^{k_n}(1+itX_{n,j})$$

$$\times\exp\left\{-\frac{t^2}{2}\sum_{j=1}^{k_n}X_{n,j}^2+\sum_{j=1}^{k_n}\gamma(tX_{n,j})\right\}$$

由式(2.10.2)并令

$$U_n\triangleq\exp\left\{-\frac{t^2}{2}\sum_{j=1}^{k_n}X_{n,j}^2+\sum_{j=1}^{k_n}\gamma(tX_{n,j})\right\}\quad(2.10.3)$$

所以

$$I_n=T_ne^{-\frac{t^2}{2}}+T_n(U_n-e^{-\frac{t^2}{2}})$$

由绝对值不等式，有

$$\left|\sum_{j=1}^{k_n}\gamma(tX_{n,j})\right|\leqslant\sum_{j=1}^{k_n}|\gamma(tX_{n,j})|$$

$$\leqslant K\Sigma|tX_{n,j}|^3\leqslant K|t|^3\max_{j\leqslant k_n}|X_{n,j}|\sum_{j=1}^{k_n}|X_{n,j}|^2\quad(2.10.4)$$

这里K为常数。根据定理条件(3)，(4)和式(2.10.4)，可知

$$\sum_{j=1}^{k_n}\gamma(t,X_{n,j})\xrightarrow[n\to\infty]{P}0$$

由式(2.10.3)和条件(3)以及依测度收敛的性质可知

$$U_n\xrightarrow[n\to\infty]{P}e^{-\frac{t^2}{2}}$$

由于$\{T_n\}$一致可积[条件(2)]，$E|I_n|^2=1$，因此$\{I_n\}$也一致可积且

$$T_n(U_n-e^{-\frac{t^2}{2}})=I_n-T_n\exp\left(-\frac{t^2}{2}\right)$$

也一致可积。

于是[2]

$$EI_n-ET_n\exp\left(-\frac{t^2}{2}\right)\xrightarrow[n\to\infty]{}0$$

即

$$El_n \xrightarrow[n\to\infty]{} e^{-\frac{t^2}{2}}$$
∎

如果 $X_{n,i}$ 关于 $\mathscr{F}_{n,i}$ 可测且

$$E(X_{n,j}|\mathscr{F}_{n,j-1})=0,\quad \forall n\geqslant 1,\ j\geqslant 1$$

则称 $\{X_{n,j},\ 1\leqslant j\leqslant k_n\}$ 为双下标鞅差序列。

定理 2.10.2 （McLeish 定理） 设 $\{X_{n,j}:\ 1\leqslant j\leqslant k_n\}$ 是鞅差序列，它满足下列条件：

(1) $\max\limits_{j\leqslant k_n} EX_{n,j}^2$ 一致有界；

(2) $\max\limits_{j\leqslant k_n}|X_{n,j}| \xrightarrow[n\to\infty]{P} 0$；

(3) $\sum\limits_{j=1}^{k_n} X_{n,j}^2 \xrightarrow[n\to\infty]{P} 1$。

那么式 (2.10.1) 的 S_n 分布收敛于 $N(0,1)$ 分布。

证明 本定理的条件 (2)，(3) 即为定理 2.10.1 的条件 (4)，(3)。如能证明在本定理条件下，$X_{n,j}$ 也满足定理 2.10.1 的条件 (1)，(2)，那么本定理就被证明了。

为此令

$$Z_{n,j}=X_{n,j}\chi\left(\sum_{k=1}^{j-1}X_{n,k}^2\leqslant 2\right) \tag{2.10.5}$$

这里 $\chi(\cdot)$ 表示集的示性函数。

因为可测集的示性函数必可测，而且由引理 2.4.4 可得

$$E(Z_{n,j}|\mathscr{F}_{n,j-1})=E\left(X_{n,j}\chi\left[\sum_{k=1}^{j-1}X_{n,k}^2\leqslant 2\right]\Big|\mathscr{F}_{n,j-1}\right)$$

$$=\chi\left[\sum_{k=1}^{j-1}X_{n,k}^2\leqslant 2\right]E(X_{n,j}|\mathscr{F}_{n,j-1})=0$$

由于 $\{X_{n,j}\}$ 是鞅差序列，因此 $\{Z_{n,j}:1\leqslant j\leqslant k_n\}$ 也是鞅差序列。

又因为

$$P[Z_{n,i} \neq X_{n,i}, \forall \text{某一个 } j \leqslant k_n]$$

$$\leqslant P\left(\sum_{j=1}^{k_n} X_{n,j}^2 \geqslant 2\right) \xrightarrow[n \to \infty]{} 0$$

即对充分大的 n 有

$$Z_{n,j} = X_{n,j}, \text{ a.s.}$$

因此对于充分大的 n, $\{Z_{n,j}\}$ 满足定理 2.10.2 的所有条件。

定义

$$T_n = \prod_{j=1}^{k_n} (1 + itX_{n,j})$$

因为

$$ET_n = E\left\{\prod_{j=1}^{k_{n-1}} (1 + itX_{n,j}) E(1 + itX_{n,k_n} | \mathscr{F}_{n,k_n-1})\right\}$$

$$= E\prod_{j=1}^{k_{n-1}} (1 + itX_{n,j})$$

$$= E\left\{\prod_{j=1}^{k_{n-2}} (1 + itX_{n,j}) E(1 + itX_{n,k_{n-1}} | \mathscr{F}_{n,k_{n-2}})\right\}$$

$$= E\prod_{j=1}^{k_{n-2}} (1 + itX_{n,j}) = \cdots = 1$$

所以定理 2.10.1 的条件 (1) 成立。

定义随机变量:

$$J_n = \begin{cases} \min\left\{j \leqslant k_n: \sum_{i=1}^{j} X_{n,i}^2 > 2\right\}, & \text{当} \sum_{j=1}^{k_n} X_{n,j} > 2 \\ k_n, & \text{其他} \end{cases} \tag{2.10.6}$$

根据式 (2.10.5) 和 (2.10.6) 可得

$$Z_{n,j} = \begin{cases} X_{n,j}, & \text{当 } 1 \leqslant j \leqslant J_{n-1} \\ X_{n,J_n}, & \text{当 } j = J_n \\ 0, & \text{当} j > J_n \end{cases}$$

因此

$$
\begin{aligned}
E|T_n|^2 &= E\prod_{j=1}^{k_n}(1+t^2Z_{n,j}^2) \\
&= E\prod_{j=1}^{J_n}(1+t^2X_{n,j}^2) \\
&\leqslant E(1+t^2X_{n,J_n}^2)\exp\left(t^2\sum_{j=1}^{J_n-1}X_{n,j}^2\right) \\
&\leqslant e^{2t^2}(1+t^2EX_{n,J_n}^2)
\end{aligned}
$$

由定理 2.10.2 条件 (1)，可知 $\{T_n\}$ 一致可积，即定理 2.10.1 的条件 (2) 成立。 ∎

2.11 连续时间情形下的鞅

提要 本节讨论连续时间上鞅的有关不等式及收敛定理。对于连续时间(上)鞅，我们可以得到与离散时间情况相类似的一系列结果。

给定一个滤波空间 $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$。随机过程 $\{X_t\}_{t\in R_+}$，$\mathscr{F}_t\subset\mathscr{F}_s$，$\forall t<s$，如果对一切 $t\in R_+$，X_t 为 $\mathscr{F}_t$ 可测，则称 $\{X_t\}$ 为 $\mathscr{F}_t$ 适应的。

在连续时间情况下，我们可以得到与离散时间情况相类似的一系列结果。设 $\{X_t\}_{t\in R_+}$ 为一适应过程，$u=\{t_1, t_2, \cdots, t_n\}$ 为 R_+ 上的一个有限子集(不妨设 u 中的元素已按大小次序排列)。我们用 $N(a, b, X, n)$ 表示序列 $\{X_{t_1}, \cdots, X_{t_n}\}$ 上穿区间 $[a, b]$ 的次数。

对 R_+ 的任何子集 D，我们令

$$N(a, b, X, D)=\sup_{\substack{u\text{为有限集}\\ u\subset D}}N(a, b, X, u)$$

设 $D=\{t_1, t_2, \cdots\}$ 为可数集，令 u_n 为 $\{t_1, t_2, \cdots, t_n\}$ 按大小重新排列后的集合，显然有

$$N(a, b, X, D)=\lim_{n\to\infty} N(a, b, X, u_n).$$

定理 2.11.1 设 $\{X_t\}_{t\in R_+}$ 为上鞅，D 为 R_+ 的可数稠密集，对任何 $r<s\in R_+$，任何 $a<b\in R$ 及 $\lambda>0$，我们有

$$E[N(a, b, X, D)\cap[r, s]]\leqslant\frac{1}{b-a}E(X_s-a)^- \tag{2.11.1}$$

如果 $\{X_t\}$ 的几乎所有轨道 $X_t(\omega)$ 为 R_+ 的右连续函数，则上述不等式中的 $D\cap[r,s]$ 可用 $[r,s]$ 代替.

证明 设 $D\cap[r, s]=\{t_1, t_2, \cdots\}$，令 $u_n=\{t_1, \cdots, t_n\}$，则用 u_n 代替 $D\cap[r, s]$ 后可由定理2.6.2推得式(2.11.1). 由于 $\{X_t-a\}$ 是上鞅，故 $\{(X_t-a)\wedge 0\}$ 也是上鞅. 也可证 $[X_t-a>0]\in\mathscr{F}_t$，设 $t<s$，对任何 $A\in\mathscr{F}_t$，有

$$\begin{aligned}
&\int_A[(X_t-a)\wedge 0]dP\\
&=\int_{A\cap[X_t-a\leqslant 0]}(X_t-a)dP\\
&\geqslant\int_{A\cap[X_t-a\leqslant 0]}(X_s-a)dP\\
&\geqslant\int_{A\cap[X_t-a\leqslant 0]}[(X_s-a)\wedge 0]dP\\
&\geqslant\int_A\{(X_s-a)\wedge 0\}dP
\end{aligned}$$

于是

$$(X_t-a)^-=-(X_t-a)\wedge 0 \text{ 是下鞅}$$

$$E(X_s-a)^-\geqslant E(X_t-a)^-$$

在

$$E[N(a, b, X, u_n)] \leqslant \frac{1}{b-a} E(X_s - a)^-$$

中令 $n \to \infty$，即得

$$E[N(a, b, X, D \cap [r, s])] \leqslant \frac{1}{b-a} E(X_s - a)^-$$

定理的最后断言显然。 ∎

下面的 Doob 不等式直接由定理 2.6.3 推出。

定理 2.11.2 设 $\{X_t\}_{t \in R_+}$ 是非负下鞅，其几乎所有的轨道右连续。令

$$X^* = \sup_t X_t$$

则

$$\|X^*\|_p \leqslant q \sup_t \|X_t\|_p \tag{2.11.2}$$

其中 $p > 1$，$q > 1$ 为一对共轭指数，即

$$\frac{1}{p} + \frac{1}{q} = 1$$

下面讨论连续时间情形下的上鞅收敛定理。

定理 2.11.3 设 $\{X_t\}_{t \in R_+}$ 是上鞅，D 为 R_+ 的可数稠密子集，则对几乎所有的 ω，$\forall\ t \in R_+$

$$\lim_{\substack{s \in D \\ s \downarrow t}} X_s(\omega) = X_{t^+}(\omega) \tag{2.11.3}$$

存在且有限。如果 $\{X_t\}$ 的几乎所有轨道为右连续，则对几乎所有的 ω，$\forall\ t \in R_+ - \{0\}$

$$\lim_{\substack{s \in R_+ \\ s \uparrow t}} X_s(\omega) = X_{t^-}(\omega) \tag{2.11.4}$$

存在且有限。

证明 设 $t \in R_+$，$a < b \in R$，令

$$H_{t,a,b} = \{\omega \mid \sup_{s \in D \cap [0,t]} |X_s(\omega)| = \infty \text{ 或 } N(a, b, X, D \cap [0, t]) = \infty\}$$

则 $H_{t,a,b}\in\mathscr{F}_t$。由定理 2.11.1 及 2.11.2 知，$P(H_{t,a,b})=0$。

我们用 Q 表示 R 中的有理数全体，令

$$H_t=\bigcup_{a<b\in Q}H_{t,a,b},\quad H=\bigcup_{t\in R_+}H_t=\bigcup_{n=1}^{\infty}H_n$$

则 $H_t\in\mathscr{F}_t$，$H_t\uparrow H$，且 $P(H)=0$。

若 $\omega\notin H$，则 $\forall\ t\in R_+$（相应地 $t\in R_+-\{0\}$），$\lim\limits_{s\in D,s\downarrow t}X_s(\omega)$（相应地 $\lim\limits_{s\in D,s\uparrow t}X_s(\omega)$）存在且有限。

如果 $\{X_t\}_{t\in R_+}$ 的几乎所有的轨道右连续，则在以上证明中可用 $[0,t]$ 代替 $D\cap[0,t]$。于是，当 $\omega\in H$ 时，$\forall$ $t\in R_+-\{0\}$，$\lim\limits_{s\in R_+,s\uparrow t}X_s(\omega)$ 存在且有限。

定理 2.11.4 设 $\{X_t\}_{t\in R_+}$ 为上鞅，D 为 R_+ 中的一个可数稠密集，则存在一个 $\{\mathscr{F}_t\}$ 适应过程 $\{\bar{X}_t\}$ 使得

(1) $\{\bar{X}_t\}$ 所有的轨道右连续，且对几乎所有 ω，$\forall t\in R_+$，有

$$\bar{X}_t=\lim_{s\in D,s\downarrow t}X_s(\omega)\tag{2.11.5}$$

(2) 对几乎所有 ω，$\forall t\in R_+-\{0\}$，有

$$\bar{X}_{t^-}(\omega)=\lim_{s\in R_+,s\uparrow t}\bar{X}_s(\omega)$$

存在且有限，此外还有

$$\bar{X}_{t^-}(\omega)=\lim_{s\in D,s\uparrow t}X_s(\omega)\tag{2.11.6}$$

(3) 对 $s>t\in R_+$，有

$$\bar{X}_t\geqslant E[\bar{X}_s|\mathscr{F}_t]\tag{2.11.7}$$

即 $(\bar{X}_t,\mathscr{F}_t)$ 为上鞅。

证明 我们沿用定理 2.11.3 证明中的记号。对 $t\in R_+$，令 $H_{t^+}=\bigcap\limits_{s>t}H_s$ 则 $H_{t^+}\in\mathscr{F}_{t^+}=\mathscr{F}_t$。如果 $\omega\notin H_{t^+}$，则存在 $t_1>t$，使 $\omega\notin H_{t_1}$，故极限 $\lim\limits_{s\in D,s\downarrow t}X_s(\omega)$ 存在且有

限。令

$$\overline{X}_t(\omega)=\begin{cases}\lim\limits_{s\in D, s\downarrow t} X_s(\omega), & \omega \notin H_{t^+}\\ 0, & \omega\in H_{t^+}\end{cases} \tag{2.11.8}$$

显然 $\{\overline{X}_t\}$ 为 $\{\mathscr{F}_t\}$ 适应过程。

证明 $\{\overline{X}_t\}$满足所要求的性质：

(1) 设 $t\in R_+$，$\omega\in H_{t^+}$，则$\forall$ $s>t$，$\omega\in H_{s^+}$。于是当 $s\geqslant t$ 时，$\overline{X}_s(\omega)=0$，从而使 $\overline{X}_\cdot(\omega)$在 t 处右连续。设$\omega\notin H_{t^+}$，由于

$$H_{t^+}=\bigcap_{r>t} H_{r^+}$$

于是存在 $r_0>t$，使$\forall$ $r_0\geqslant r>t$，且 $\omega\notin H_{r^+}$。给定 $\varepsilon>0$，取 $\delta>0$，$\delta<r_0-t$，使当 $s\in D$，$s>t$，且 $s-t<\delta$时，$|\overline{X}_t(\omega)-X_s(\omega)|\leqslant\varepsilon$，这样，当 $r>t$，且 $r-t<\delta$ 时，就有

$$|\overline{X}_t(\omega)-\overline{X}_r(\omega)|=\lim_{s\in D, s\downarrow r}|\overline{X}_t(\omega)-X_s(\omega)|\leqslant\varepsilon$$

这表明 $\overline{X}_\cdot(\omega)$ 在 t 处右连续。因此$\overline{X}_t$的一切轨道在 R_+ 上右连续。最后，如果 $\omega\notin H$，则$\forall$ $t\in R_+$，由式(2.11.8)得式(2.11.5)。

(2) 设 $t>0$，$\omega\notin H$，则 $\lim\limits_{s\in D, s\uparrow t} X_s(\omega)$ 存在且有穷。与(1)类似地可证明式(2.11.6)。

(3) 设 $s>t\in R_+$，令 $t_n\in D$，$t_n<s$ 且 $t_n\downarrow t$。又令 $S_n\in D$，$S_n\downarrow s$，则对任何 $A\in\mathscr{F}_t$，我们有

$$\int_A X_{t_n}dP\geqslant\int_A X_{s_n}dP$$

由于$\{X_{t_n}\}$及$\{X_{s_n}\}$一致可积（定理 2.7.4），在上式两边令 $n\to\infty$ 得

$$\int_A \overline{X}_t dP\geqslant\int_A \overline{X}_s dP$$

这表明 $\{\bar{X}_t, \mathscr{F}_t\}$ 是上鞅。

设 $\{X_t\}_{t\in R_+}$, $\{Y_t\}_{t\in R_+}$ 为两个随机过程，称 $\{X_t\}$ 为 $\{Y_t\}$的修正，如果对一切 $t\in R_+$, $X_t = Y_t$ a.s.; 如果对几乎所有 ω, 轨道 $X_\cdot(\omega)$ 与 $Y_\cdot(\omega)$ 一致，则称 $\{X_t\}$ 与 $\{Y_t\}$ 无区别。

定理 2.11.5 设 $\{X_t\}_{t\in R_+}$ 为 $\{\mathscr{F}_t\}$ 上鞅 ($(X_t, \mathscr{F}_t)_{t\in R_+}$ 为上鞅)，则为要 $\{X_t\}$ 有右连续适应修正，必须而且只需 R_+ 上的函数：$t\to E[X_t]$ 为右连续。

证明 任取 R_+ 的一个可数稠密集 D. 从定理 2.11.4 得到相应的上鞅 $\{\bar{X}_t\}$, 其中

$$\bar{X}_t = \lim_{s\in D, s\downarrow t} X_s(\omega), \text{ a.s.}$$

设 $r_n\in D$, $r_n\downarrow t$, 对任何 $A\in\mathscr{F}_t$, 我们有

$$\int_A X_t dP \geqslant \int_A X_{r_n} dP$$

由于 $\{X_{r_n}\}$ 一致可积，从而 $X_{r_n}\xrightarrow{L_1}\bar{X}_t$, 故

$$\int_A X_t dP \geqslant \int_A \bar{X}_t dP$$

此即 $X_t \geqslant E(\bar{X}_t|\mathscr{F}_t)$, a.s..

由于两者都 $\mathscr{F}_t$ 可测，因而 $X_t\geqslant\bar{X}_t$, a.s..

为要 $X_t = \bar{X}_t$, a.s., 或者等价地 $EX_t = E\bar{X}_t$ (因为 $X_t\geqslant\bar{X}_t$ a.s.), 必须且只需

$$EX_t = \lim_{n\to\infty} EX_{r_n}$$

由于 $s\to EX_s$ 是单调非增函数，这等价于它在 t 处右连续. 因此，如果函数 $s\to EX_s$ 在 R_+ 上右连续，则 $\{\bar{X}_t\}$ 为 $\{X_t\}$ 的右连续适应修正。

反之，若存在 $\{X_t\}$ 的右连续修正 $\{Y_t\}$，则 $EX_t = EY_t$，根据上述论证，$t \to EY_t = EX_t$ 为 R_+ 上的右连续函数。 ∎

关于连续时间情形下的鞅收敛定理，其证明与离散时间情况相似，所以我们只叙述结果。

定理 2.11.6 设 $\{X_t\}_{t \in R_+}$ 为上鞅，其几乎所有轨道右连续。如果

$$\sup_t EX_t^- < \infty$$

则当 $t \to \infty$ 时，X_t a.s. 收敛于一个可积随机变量 X_∞。

定理 2.11.7 设 $\{X_t\}_{t \in R_+}$ 是一致可积上鞅，其几乎所有轨道右连续，则当 $t \to \infty$ 时

$$X_t \xrightarrow{\text{a.s.}, L_1} X_\infty$$

并且 X_∞ 可积。

2.12 鞅型序列及其性质[1]

提要 拟鞅、渐近鞅、依概率渐近鞅、极限鞅、依概率极限鞅、循序鞅、终鞅、拟终鞅、L_1 极限鞅等均可看作鞅的拓广或一般化。在本节中，我们把这些鞅的拓广通称为鞅型序列。本节给出鞅型序列的定义，并讨论各种鞅型序列的性质。

鞅型序列是在实际应用中提出来的，它与鞅具有类似的或部分类似的性质。本世纪 60 年代中期以来，讨论了下列各种鞅型序列：

设 $(\Omega, \mathscr{F}, P)$ 是一完备概率空间，$\mathscr{F}_n$ 是 $\mathscr{F}$ 的上升

1）本节主要根据参考文献［15，22］和汪振鹏同志提供的书面材料写成。

子σ代数序列，T是$(\mathscr{F}_n)_{n\geq 1}$有界停时的全体。

对于 $\sigma\in T$，定义

$$T(\sigma)\triangleq\{\tau|\tau\in T,\ \tau\geqslant\sigma\}$$

下面讨论的$(X_n,n\geqslant 1)$均为$(\Omega,\mathscr{F},P)$上与$\mathscr{F}_n$适应的随机变量序列。

定义 2.12.1 若 $-\infty<\lim\limits_{\tau\in T}EX_\tau<\infty$，即对于任给的 $\varepsilon>0$，存在 $\delta>0$，当 $\tau>\delta$ 时

$$|EX_\tau-\lim_{\tau\in T}EX_\tau|<\varepsilon$$

则称$(X_n,n\in \boldsymbol{N})$为渐近鞅，简记为 amart。

定义 2.12.2 若

$$\sum_{n=1}^{\infty}E|E(X_{n+1}|\mathscr{F}_n)-X_n|<\infty$$

则称$(X_n,n\geqslant \boldsymbol{N})$为拟鞅，简记为 QM。

定义 2.12.3 若对任给的 $\varepsilon>0$

$$\lim_{\sigma\in T}\sup_{\tau\in T(\sigma)}P\{|E(X_\tau|\mathscr{F}_\sigma)-X_\sigma|>\varepsilon\}=0$$

则称$(X_n,n\in \boldsymbol{N})$为依概率渐近鞅，简记为 pramart。

定义 2.12.4 若

$$\lim_n\sup_{m\geqslant n}|E(X_m|\mathscr{F}_n)-X_n|=0,\ \text{a.s.}$$

则称$(X_n,n\in \boldsymbol{N})$为 I 型极限鞅，简记为 mil(1)。

定义 2.12.5 若对任给的 $\varepsilon>0$

$$\lim_{\sigma\in T}P\{\sup_{m\geqslant\sigma}|E(X_m|\mathscr{F}_\sigma)-X_\sigma|>\varepsilon\}=0$$

则称$(X_n,n\in \boldsymbol{N})$为 II 型极限鞅，简记为 mil(2)。

定义 2.12.6 若对任给的 $\varepsilon>0$

$$\lim_{\sigma\in T}\sup_{m\geqslant\sigma}P\{|E(X_m|\mathscr{F}_\sigma)-X_\sigma|>\varepsilon\}=0$$

则称$(X_n,n\in \boldsymbol{N})$为 III 型极限鞅，简记为 mil(3)。

定义 2.12.7 若对任给的 $\varepsilon>0$

$$\lim_{n} P\ \{\sup_{m\geq n}|E(X_m|\mathscr{F}_n)-X_n|>\varepsilon\}=0$$

则称 $(X_n, n\in N)$ 为依概率极限鞅(1)，简记为 GFT(1)。

定义 2.12.8 若对任给的 $\varepsilon>0$

$$\lim_{n}\sup_{m\geq n} P\{|E(X_m|\mathscr{F}_n)-X_n|>\varepsilon\}=0$$

则称 $(X_n, n\in N)$ 为依概率极限鞅 (2)，简记为 GFT(2)。

定义 2.12.9 若存在上升适应集合序列 $(A_n)_{n\geq1}$，这里 $A_n\in\mathscr{F}_n$，$A_n\subset A_{n+1}$，$n=1,2,\cdots$，且有

$$\bigcup_{n=1}^{\infty} A_n=\Omega \text{ 和 } E(X_{n+1}|\mathscr{F}_n)=X_n,\ \omega\in A_n,\ \forall n\geq1$$

则称 $(X_n, n\geq1)$ 为循序鞅，简记为 PM。

定义 2.12.9 也可换一种表达形式：

定义 2.12.9′ 若令 $A_n=[\omega|E(X_{n+1}|\mathscr{F}_n)=X_n]$，$A_n\subset A_{n+1}$，$\forall n\geq1$ 且

$$P\left(\bigcup_{n=1}^{\infty} A_n\right)=1$$

则称 $(X_n, n\geq1)$ 为循序鞅。

定义 2.12.10 若

$$P\{\omega: E(X_{n+1}|\mathscr{F}_n)\neq X_n,\ \text{i.o.}\}=0$$

这里"i.o."表示"无限次发生"，于是称 $(X_n, n\geq1)$ 为终鞅，简记为 EM。

定义 2.12.11 若

$$\sum_{n=1}^{\infty}|E(X_{n+1}|\mathscr{F}_n)-X_n|<\infty$$

则称 $(X_n, n\geq1)$ 为拟终鞅，简记为 QEM。

定义 2.12.12 若

$$\lim_{n} E\,[\sup_{m\geq n}|E(X_m|\mathscr{F}_n)-X_n|]=0$$

则称 $(X_n, n \geqslant 1)$ 为 I 型 L^1 极限鞅，简记为 L^1 极限鞅 (1)。

定义 2.12.13 若

$$\lim_n \sup_{m \geqslant n} E\left|E(X_m|\mathscr{F}_n) - X_n\right| = 0$$

则称 $(X_n, n \geqslant 1)$ 为 L^1 极限鞅 (2)。

下面我们不加证明地把鞅[编号为 (0)] 与 13 种鞅型序列之间的关系，用一个有向图形表示出来：

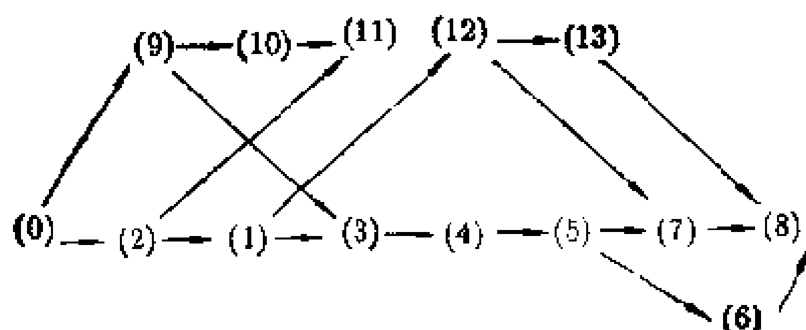


图 2.12.1 鞅与鞅型序列的关系

图中 (0) 表示鞅，(1) 表示定义 2.12.1 定义的渐近鞅，依此类推，符号 $A \Rightarrow B$ 表示"是 A 必是 B"。

下面讨论鞅型序列的 7 个性质，令 $\mathscr{X}$ 表示随机序列族。

定义 2.12.14 如果对任何停时 σ 每一个 $(X_n, n \geqslant 1) \in \mathscr{X}$，有 $(X_{\sigma \wedge n}, n \geqslant 1) \in \mathscr{X}$，则称 $\mathscr{X}$ 具有可选停时性。

Edgar 和 Sucheston (1976, 1977) 证明了渐近鞅具有可选停时性，但是依概率极限鞅与极限鞅没有这一性质。当 $\sup_T E|X_\tau| < \infty$ 时，极限鞅和依概率极限鞅也具有可选停时性。下面的定理表明循序鞅、终鞅和拟鞅也具有可选停时性。

定理 2.12.1 循序鞅、终鞅、拟鞅和拟终鞅具有可选停时性质。

证明 设 σ 为一停时。注意到，对 $n \geqslant 1$，有

$$E(X_{\sigma \wedge n}|\mathscr{F}_{n-1}) = X_\sigma \chi_{[\sigma < n]} + E(X_n|\mathscr{F}_{n-1})\chi_{[\sigma \geqslant n]} \tag{2.12.1}$$

而且由于

$$
\begin{aligned}
X_\sigma \chi_{[\sigma<n]} &= X_\sigma \chi_{[\sigma=n-1]} + X_\sigma \chi_{[\sigma<n-1]} \\
&= X_{n-1}\chi_{[\sigma=n-1]} + X_\sigma \chi_{[\sigma<n-1]} \\
&= X_{n-1}\chi_{[\sigma\geqslant n-1]} - X_{n-1}\chi_{[\sigma\geqslant n]} + X_\sigma \chi_{[\sigma<n-1]} \\
&= X_{\sigma\wedge(n-1)} - X_{n-1}\chi_{[\sigma\geqslant n]}
\end{aligned}
$$

因此式(2.12.1)可写成

$$
\begin{aligned}
E(X_{\sigma\wedge n}|\mathscr{F}_{n-1}) = X_{\sigma\wedge(n-1)} + [E(X_n|\mathscr{F}_{n-1}) \\
- X_{n-1}]\chi_{[\sigma\geqslant n]}
\end{aligned}
\tag{2.12.2}
$$

如果 $(X_n\ n\geqslant 1)$ 是循序鞅,则对于 $\omega\in A_{n-1}$ 应有

$$
E(X_n|\mathscr{F}_{n-1}) = X_{n-1}
$$

这里 $A_{n-1}\subset A_n$ 且

$$
P\left(\bigcup_{n=1}^{\infty} A_n\right) = 1
$$

对于 $\omega\in A_{n-1}$, 因为有

$$
E(X_{\sigma\wedge n}|\mathscr{F}_{n-1}) = X_{\sigma\wedge(n-1)} \tag{2.12.3}
$$

若 $\sigma<n$, 则 $\sigma<n+1$, 式(2.12.3)也成立,因此 $(X_{\sigma\wedge n}, n\geqslant 1)$ 是循序鞅,即循序鞅具有可选停时性。

由于当 $E(X_n|\mathscr{F}_{n-1}) = X_n$ 时, $E(X_{n\wedge\sigma}|\mathscr{F}_{n-1}) = X_{\sigma\wedge(n-1)}$, 故可选停时性对于终鞅成立。

最后由式(2.12.2)可得

$$
\begin{aligned}
&E|E(X_{n\wedge\sigma}|\mathscr{F}_{n-1}) - X_{\sigma\wedge(n-1)}| \\
&\leqslant E|E(X_n|\mathscr{F}_{n-1}) - X_n|
\end{aligned}
$$

如果 $(X_n, n\geqslant 1)$ 是拟鞅, 那么 $(X_{n\wedge\sigma}, n\geqslant 1)$ 也是拟鞅, 即拟鞅具有可选停时性。 ▮

设 $\mathscr{R}$ 为具有某性质的随机序列。

定义 2.12.15 设 $(X_n, n\geqslant 1)\in\mathscr{R}$, 若对任意上升趋向∞的有界停时列 τ_n, 有

$$
(X_{\tau_n}, \mathscr{F}_{\tau_n})\in\mathscr{R}
$$

则称 $\mathscr{R}$ 具有可选采样性。

这个定义意味着，若 $(X_n, n\in \boldsymbol{N})$ 为某种鞅型序列（如拟鞅），那么 $(X_{\tau_n}, \mathscr{F}_{\tau_n})$，$\tau_n\uparrow\infty$，也是某种鞅型序列（如拟鞅）。

Edgar 和 Sucheston (1976) 证明了渐近鞅具有可选采样性，Tomkins (1984) 证明了拟鞅具有可选采样性。

定理 2.12.2 拟鞅具有可选采样性。

证明 设 $(X_n, \mathscr{F}_n)_{n\geq 1}$ 是拟鞅，令 $d_n = X_n - X_{n-1}$，容易验证

$$\left\{ X_n - \sum_{j=1}^{n} E(d_j | \mathscr{F}_{j-1})\right\}_{n\geq 1}$$

是鞅，因此对任何有界停时 t 有

$$E(X_t - X_n | \mathscr{F}_n) = E\left(\sum_{j=n+1}^{t} E(d_j | \mathscr{F}_{j-1}) | \mathscr{F}_n\right), \forall t > n \tag{2.12.4}$$

现在设 $\{\tau_k\}$ 为有界停时序列，$\tau_k \to \infty$。对于 $k\geq 1$，令 $X_k^* = X_{\tau_k}$，$\mathscr{G}_k = \mathscr{F}_{\tau_k}$。

对于 $n\geq 1$，由式 (2.12.4) 可得

$$\begin{aligned}
&E(\chi_{[\tau_{k-1}=n]} | E(X_k^* | \mathscr{G}_{k-1}) - X_{k-1}^* |) \\
&= E\left(\chi_{[\tau_{k-1}=n]} \left| E\left(\sum_{j=n+1}^{k} E(d_j | \mathscr{F}_{j-1}) | \mathscr{F}_n\right)\right|\right) \\
&\leq E\left(E\left(\chi_{[\tau_{k-1}=n]} \sum_{j=n+1}^{\tau_k} | E(d_j | \mathscr{F}_{j-1}) | \mathscr{F}_n\right)\right) \\
&= E\left(\chi_{[\tau_{k-1}=n]} \sum_{j=\tau_{k-1}+1}^{\tau_k} \left| E(d_j | \mathscr{F}_{j-1}) \right|\right)
\end{aligned}$$

不等式两边对 $n\geq 1$ 取和，对 $k\geq 1$ 可得

$$E | E(X_k^* | \mathscr{G}_{k-1}) - X_{k-1}^* | \leq E\left(\sum_{j=\tau_{k-1}+1}^{\tau_k} | E(d_j | \mathscr{F}_{j-1}) |\right)$$

所以

$$\sum_{k=2}^{\infty} E|E(X_k^*|\mathscr{G}_{k-1}) - X_{k-1}^*|$$

$$\leqslant \sum_{k=2}^{\infty} E\left(\sum_{j=\tau_{k-1}+1}^{\tau_k} |E(d_j|\mathscr{F}_{j-1})|\right)$$

$$\leqslant \sum_{n=1}^{\infty} E|(E(d_n|\mathscr{F}_{n-1}))| < \infty$$

即 $\{X_k^*\}$ 是拟鞅。 ▮

定义 2.12.16 设 $(X_n, \mathscr{F}_n) \in \mathscr{X}$，若 $X_n = Y_n + Z_n$，$\forall n \geqslant 1$，$(Y_n, \mathscr{F}_n)$ 为鞅，且

$$\lim_n EX_A Z_n = 0, \forall A \in \bigcup_n \mathscr{F}_n$$

于是称 $\mathscr{X}$ 有 Riesz 分解性。

可以证明上(下)鞅具有 Riesz 分解性 (见 2.9 节)，渐近鞅与 L^1 极限鞅 (1) 也具有 Riesz 分解性，但依概率渐近鞅不具有 Riesz 分解性。

注意：$\{X_n\}$ 具有 Riesz 分解性的必要条件是 $E(X_n)$ 收敛。

事实上，若 $X_n = Y_n + Z_n$，$(Y_n, n \in \boldsymbol{N})$ 为鞅，且

$$EX_A Z_n \to 0, \forall A \in \bigcup_n \mathscr{F}_n$$

就有

$$E(X_n) = E(Y_n) + E(Z_n X_\Omega) = E(Y_1) + E(Z_n X_\Omega) \to E(Y_1)$$

定义 2.12.17 设 $(X_n, \mathscr{F}_n) \in \mathscr{X}$，若

$$\sup_n E|X_n| < \infty$$

有

$$\sup_\lambda[\lambda P\{\sup_n |X_n| > \lambda\}] < \infty$$

则称 $\mathscr{X}$ 具有极值不等式性质。

Edgar 和 Sucheston (1967, 1977 年)已经证明拟鞅和渐近鞅具有极值不等式性质,极限鞅和 GFT 不具有此性质。

定义 2.12.18 设 $(X_n, \mathscr{F}_n) \in \mathscr{X}$, 若

$$\sup_n E|X_n| < \infty$$

则

$$\lim_n X_n, \text{ a.s.}$$

存在,称 $\mathscr{X}$ 有 a.s. 收敛性.例如鞅具有 a.s. 收敛性。

定义 2.12.19 设 $(X_n, \mathscr{F}_n) \in \mathscr{X}$, 若对任一 $(\mathscr{F}_n)$ 可料序列 $\{\xi_n\}$, 有

(1) $(\eta_n, \mathscr{F}_n) \in \mathscr{X}$, 其中

$$\eta_n = \sum_{k=1}^{n} \xi_k (X_k - X_{k-1})$$

(2) 若 $\sup_n E|X_n| < \infty$, 则 $\{\eta_n\}$ 在

$$\{\sup_n |\xi_n| < \infty\}$$

上 a.s. 收敛。这一性质称为变换性。

定义 2.12.20 设 $(X_n, \mathscr{F}_n) \in \mathscr{X}$, 若

$$\sup_n E|X_n| < \infty$$

则

$$\sum_{n=1}^{\infty} (X_n - X_{n-1})^2 < \infty \text{ a.s.}$$

这一性质称为平方可和性。

定义 2.12.21 设 $\mathscr{Y} = \{(X_n, \mathscr{F}_n), (X_n, \mathscr{F}_n) \in \mathscr{X}, \sup_n E|X_n| < \infty\}$, 若 $\mathscr{Y}$ 是向量格,则称 $\mathscr{X}$ 有格性。

我们引进如下记号: A 表示可选停时性; B 表示可选采样性; C 表示 Riesz 分解性; D 为极值不等式性; E 为 a.s. 收敛性; F 为变换性; G 为平方可和性; H 为格性。

这样，鞅和鞅型序列的性质可以用表 2.12.1 表示：

表 2.12.1 鞅型序列的性质

序列＼性质	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*
鞅	✓	✓	✓	✓	✓	✓	✓	×
拟 鞅	✓	✓	✓	✓	✓	×	✓	✓
渐近鞅	✓	✓	✓	✓	✓	×	×	✓
依概率渐近鞅	✓	✓	⊗	×△	✓	×	×	✓
极限鞅	×△	×△	⊗	×△	✓	×	×	×△
依概率极限鞅	×△	×△	⊗	×△	×	×	×	×△
循序鞅	✓	×	⊗	×△	✓	✓	✓	×
终 鞅	✓	×	⊗	×△	⊗	⊗	⊗	×

注：“✓”表示有此性质；

“×”表示没有此性质；

“×△”表示当 $\sup_{\tau} E|X_\tau|<\infty$ 时有此性质；

“⊗”表示当 $\sup_n |X_n| \in L_1$ 时有此性质.

2.13 鞅型序列的收敛定理

提要 本节不加证明地介绍鞅型序列的收敛定理，具体证明参见所注文献.

在应用中出现的随机序列不一定恰好是鞅，可能是某种具有鞅的某些性质的序列（例如在自适应控制的渐近分析中遇到的几乎上鞅序列，见 4.2 节），因此，我们在本节中不加证明地介绍一些鞅型序列的收敛定理是有好处的. 读者如果遇到这方面的问题，可以在参考文献中找到更详尽的陈述和证明.

本节所用的概念和记号在上一节均有详细说明。

定理 2.13.1 设 $(X_n, \mathscr{F}_n)$ 是 mil(3)，若

$$\liminf_{n\to\infty} E|X_n| < \infty$$

则 X_n a.s. 收敛且

$$|\lim_n X_n| < \infty$$

定理的证明参见文献[19]。

定理 2.13.2 设 $(X_n, \mathscr{F}_n)$ 是实值 GFT(2)，若

$$\min(\lim_n \inf EX_n^+, \lim_n \inf EX_n^-) < \infty$$

则 $X_n \xrightarrow{P} X$，且 $|X| < \infty$。

证明见文献 [24]。

定理 2.13.3 若 $\{X_\tau\}_{\tau\in T}$ 一致可积，其中 T 是全体有界停时，那么下述命题等价：

(1) $(X_n, n\in \boldsymbol{N})$ a.s. 收敛；

(2) $(X_n, \mathscr{F}_n)$ 是 amart。

证明见文献 [25]。

定理 2.13.4 若 $\sup_T E|X_\tau| < \infty$，则下述命题等价：

(1) $(X_n, n\in \boldsymbol{N})$ a.s. 收敛；

(2) $(X_n, \mathscr{F}_n)$ 是 pramart。

证明见文献 [20]。

定理 2.13.5 若 $(X_n, n\in \boldsymbol{N})$ 一致可积，则下述命题等价：

(1) $(X_n, n\in \boldsymbol{N})$ a.s. 收敛；

(2) $(X_n, n\in \boldsymbol{N})$ 是 mil(3)。

证明见文献 [19]。

设 $\bar{T}$ 是全体 $(\mathscr{F}_n, n\in \boldsymbol{N})$ 停时。

定理 2.13.6 设 $(X_n, \mathscr{F}_n)$ 是 QEM，若

$$E|X_\tau|\chi_{[\tau<\infty]}<\infty,\ \forall\tau\in\bar{T}$$

那么$(X_n, n\in \boldsymbol{N})$ a.s. 收敛且$|\lim_n X_n|<\infty$。

证明见文献[26]。

定义 2.13.1 若 $\forall\varepsilon>0$，有

$$\lim_{\sigma\in T}\ \sup_{\tau\in T(\sigma)} P\{[E(X_\tau|\mathscr{F}_\sigma)-X_\sigma]^->\varepsilon\}=0$$

则称$(X_n, \mathscr{F}_n)$为 subpramart（依概率渐近下鞅）。

定理 2.13.7 设$(X_n, \mathscr{F}_n)$是 subpramart，若

$$\lim_n\inf EX_n^+<\infty$$

则$(X_n)_{n\geqslant1}$ a.s. 收敛。

证明见文献[18]。

定义 2.13.2 若

$$\sum_{k=1}^{\infty}[E(X_{n+1}|\mathscr{F}_n)-X_n]^-<\infty,\ \text{a.s.}$$

则称$(X_n, \mathscr{F}_n)$是 QEM$^-$。

定理 2.13.8 设$(X_n, \mathscr{F}_n)$是 QEM$^-$，若

$$EX_\tau^+\chi_{[\tau<\infty]}<\infty,\ \forall\tau\in\bar{T}$$

那么$(X_n)_{n\geqslant1}$ a.s. 收敛，且

$$\lim_n X_n>-\infty,\ \text{a.s.}$$

证明 令

$$Y_1=0,\ Y_n=\sum_{i=1}^{n-1}[E(X_{i+1}|\mathscr{F}_i)-X_i]^-,\ n\geqslant2$$

并设

$$Z_n=X_n+Y_n,\ n\geqslant1$$

则$(Z_n, \mathscr{F}_n)$是下鞅。对于某一取定的实数$a>0$，令

$$\rho=\inf\{n: Y_{n+1}^+>a\},\ \inf\varnothing=\infty$$

则$\rho\in\bar{T}$，令$\tau_n=\rho\wedge n,\ n\geqslant1$ 则

$$(\tau_n)_{n\geq 1}\subset T$$

记

$$u_n = Z_{\tau_n},\ \mathscr{G}_n = \mathscr{F}_{\tau_n},\ n\geq 1$$

由定理 2.5.1 知 $(u_n, \mathscr{G}_n)$ 是下鞅，记 $(\mathscr{G}_n)_{n\geq 1}$ 停时全体为 $\overline{T}(\mathscr{G}_n)$，对任意的 $\sigma\in\overline{T}(\mathscr{G}_n)$，令

$$\tau = \begin{cases}\tau_k, & \omega\in(\sigma = k),\ k = 1, 2, \cdots\\ \rho, & \omega\in(\sigma = \infty)\end{cases}$$

则 $\tau = \rho\wedge\sigma\in\overline{T}$，有

$$\int_{(\sigma<\infty)} u_\sigma^+ dP = \sum_{k=1}^{\infty}\int_{(\sigma=k)} u_k^+ dP$$

$$\leqslant \sum_{k=1}^{\infty}\int_{\sigma=k}(X_{\rho\wedge k}^+ + Y_{\rho\wedge k}^+)dP$$

$$\leqslant a + \sum_{k=1}^{\infty}\int_{\rho=k} X_{\rho\wedge k}^+ dP$$

$$\leqslant a + \int_{(\tau<\infty)} X_\tau^+ dP < \infty$$

由此可知 $(u_n)_{n\geq 1}$ a.s. 收敛且

$$\lim_n u_n > -\infty^{1)}$$

因而在 $\rho=\infty$ 上 (Z_n) a.s. 收敛，而且

$$(\sup_n Y_n\leqslant a)\subset(\lim X_n = X)$$

由 a 的任意性知 $(X_n)_{n\geq 1}$ a.s. 收敛且

$$\lim_n X_n > -\infty,\ \text{a.s.}\qquad\blacksquare$$

我们把适应可积序列 $(X_n, \mathscr{F}_n)$ 满足的条件用下列记号来表示：

(D) $\sup_n E|X_n| < \infty$；$(\mathrm{D}^{-(+)})$ $\sup_n EX_n^{-(+)} < \infty$

---

1) 见文献[21].

(B) $\sup_{T} E|X_\tau| < \infty$; ($B^{-(+)}$) $\sup_{T} EX_\tau^{-(+)} < \infty$

(C) $E|X_\tau|\chi_{[\tau<\infty]} < \infty$, $\forall \tau \in \overline{T}$;

($C^{-(+)}$) $EX_\tau^{-(+)}\chi_{[\tau<\infty]} < \infty$, $\forall \tau \in \overline{T}$

(d) $\lim_{n} \inf E|X_n| < \infty$; ($d^{-(+)}$) $\lim_{n} \inf EX_n^{-(+)} < \infty$

(d_T) $\lim_{T} \inf E|X_\tau| < \infty$; ($d_T^{-(+)}$) $\lim_{T} \inf EX_\tau^{-(+)} < \infty$

令

$$\hat{D} = D^+ \cup D^-,\ \hat{B} = B^+ \cup B^-,\ \hat{C} = C^+ \cup C^-$$

$$\hat{d} = d^+ \cup d^-,\ \hat{d}_T = d_T^+ \cup d_T^-$$

于是有如下命题:

定理 2.13.9 设 $(X_n, \mathscr{F}_n)$ 是 amart，则下列诸条件等价:

(D^+), (D^-), (D), ($\hat{D}$), (d^+), (d^-), (d), ($\hat{d}$), (d_T^+), (d_T^-), (d_T), ($\hat{d}_T$), (B^+), (B^-), (B), (B^-), (B), ($\hat{B}$), (C)。

这些定理的条件有些还可放松。

第二章　测验题

1.设 $(X_n, \mathscr{F}_n)$ 是独立增量随机序列，

$$E\{(X_{n+1} - X_n)|\mathscr{F}_n\} = E(X_{n+1} - X_n)$$

且 $E|X_n| < \infty$，$\forall n \geqslant 1$，证明

(1) $(X_n, \mathscr{F}_n)$ 为上鞅的充分必要条件是 EX_n 单调下降.

(2) $(X_n, \mathscr{F}_n)$ 为鞅的充分必要条件是 EX_n 为常数.

2.若随机序列 $\{X_n\}$ 和 $\{Y_n\}$ 均一致可积，证明 $\{\alpha X_n + \beta Y_n\}$ 是一致可积序列，其中 α，β 为常数.

3.证明，若鞅 $(X_n, \mathscr{F}_n)$ 可料，则 $X_n = X_1$, a.s. $\forall n > 1$.

4.设 $\{X_n\}$ 是随机序列，$X_n \in \mathscr{F}_n$，$(\varepsilon_n, \mathscr{F}_n)$ 为鞅差序列，

$$E(\varepsilon_n|\mathscr{F}_{n-1}) = 0,\ E(\varepsilon_n^2|\mathscr{F}_{n-1}) = \sigma^2$$

令

$$S(n)=1+\sum_{k=1}^{n}\varphi_k^2,\ S(0)=1$$

$$Z_n=\sum_{k=1}^{n}\varphi_k e_k/S(k),\ Z_0=0$$

其中 $\varphi_k=\varphi(X_{k-1},X_{k-2},\cdots,X_0)$，$Z_n$ 均为可积纯量函数。

证明

(1) $(Z_n,\mathscr{F}_n)$ 是鞅；

(2) $Z_n\xrightarrow[n\to\infty]{}Z_\infty$，a.s. 且 Z_∞ 可积。

5.设 $(X_n,\mathscr{F}_n)$ 和 $(Y_n,\mathscr{F}_n)$ 是两个上鞅，τ 为 $\mathscr{F}_n$ 停时，$\tau\in\mathbf{Z}_+$. 在集 $\{\tau<\infty\}$上，$X_\tau\geqslant Y_\tau$.

令

$$Z_n=\begin{cases}X_n, & 当\ n<\tau\\ Y_n, & 当\ n\geqslant\tau\end{cases}$$

证明 $(Z_n,\mathscr{F}_n)$ 是上鞅。

第三章　随机积分与随机微分方程初步

随机微分方程(包括随机积分)对于控制理论的重要性是显见的：描述随机系统需要用随机微分方程，滤波与随机控制也都得从随机微分方程(或差分方程)描述的系统出发。

现举例说明：

例1　考虑一个简单的人口生长模型

$$\frac{dN}{dt}=a(t)N(t),\ N(0)=N_0 \tag{3.0.1}$$

其中 $N(t)$ 表示 t 时刻人口的多少(看作连续变量)，$a(t)$是 t 时刻的相对生长率。$a(t)$ 可能不完全清楚，它受到某些随机因素的影响，因此我们有

$$a(t)=r(t)+\text{“噪声”}$$

其中噪声的确切值未知，但其分布已给出。那么在式(3.0.1)情形下，怎样求解？

例2　设 $Q(t)$ 表示 t 时刻在电路一定点的电荷，它满足微分方程：

$$L\ddot{Q}(t)+R\dot{Q}(t)+\frac{1}{C}Q(t)=F(t),Q(0)=Q_0 \tag{3.0.2}$$

$$Q'(0)=I_0$$

其中 L 表示电感，R 表示电阻，C 表示电容，而 $F(t)$ 表示电源电动势。在某些情形下，$F(t)$ 不是完全确定的，

$$F(t)=G(t)+\text{“噪声”}$$

如何解方程(3.0.2)？

例 3（滤波问题） 为了改进解的精度，在例 2 中如果我们测得

$$Z(t_1), \cdots, Z(t_n) \tag{3.0.3}$$

$Z(t_n)$ 是 $Q(t)$ 在 $t=t_n$ 时的量测值，由于量测误差，$Z(t_n)$ 并非 $Q(t_n)$ 的确切值，即

$$Z(t_n)=Q(t_n)+\text{“噪声”} \tag{3.0.4}$$

什么是满足方程 (3.0.2) 的 $Q(t)$ 的最优估计？怎样利用测得的数据 (3.0.3) 找出 $Q(t)$ 的最优估计？如果 $t>t_n$，估计 $Q(t)$ 称为预报问题，如果 $t=t_n$，估计 $Q(t)$ 称为滤波问题。

简单的离散时间随机系统可以写成

$$x(k+1)=f(x(k), u(k), k)+e(k) \tag{3.0.5}$$

其中 $\{e(k)\}$ 是均值为零，方差为 σ^2 的独立同分布随机变量序列。对于简单的连续时间系统能否类似地写成

$$\frac{dx(t)}{dt}=f(x(t), u(t), t)+v(t) \tag{3.0.6}$$

其中 $v(t)$ 表示噪声，$Ev(t)=0$，$v(t)$，$v(s)$ 独立 $t\neq s$。

下面我们考虑 $v(t)$ 是标量的情形：

引理 3.0.1 设 $\{v(t), t\in T\}$ 是具有下列性质的连续时间的二阶矩存在的随机过程：

(1) $t\neq s$ 时 $v(t)$ 与 $v(s)$ 独立；

(2) 对所有的 $t\in T$，$v(t)$ 均方连续，且具有有限方差；

(3) $v(t)$ 的均值为零，

那么 $Ev^2(t)=0$ 即 $v(t)=0$, a.s.。

证明 因为 $v(t)$ 均方连续，所以

$$\lim_{s\to t} E[v(t)-v(s)]^2=0 \tag{3.0.7}$$

又

$$\begin{aligned} E[v(t)-v(s)]^2 &= Ev^2(t)-2Ev(t)v(s)+Ev^2(s) \\ &= Ev^2(t)+Ev^2(s) \end{aligned} \tag{3.0.8}$$

对式(3.0.8)两边取 $s \to t$ 的极限,得

$$2Ev^2(t) = \lim_{s\to t} E[v(t) - v(s)]^2 = 0$$

所以

$$Ev^2(t) = 0$$ ▮

引理 3.0.1 告诉我们连续时间的随机系统用式(3.0.6)表示是不行的,因为此时 $v(t) = 0$, a.s.. 为此,我们必须重新引入随机微分方程的解析表示式完整地描述例 1 和 2.

为了研究随机微分方程,我们首先要考察随机过程的概念,并引入马尔可夫过程的定义及有关定理. 然后讨论维纳过程的定义与性质,维纳过程的随机积分以及伊藤微分公式. 最后讨论随机微分方程解的存在性、唯一性和马氏性(马尔可夫性质的简称).

3.1 随机过程的基本概念

提要 本节介绍随机过程的存在定理,马尔可夫过程及有关定义,为今后讨论随机积分与随机微分方程作准备.

有一族随机变量 $\{X_t(\omega)\}_{t\in\mathscr{T}}$ 需要我们联合起来考虑(如连续时间情形下的鞅). 如果 $\mathscr{T}$ 是一个时间的集合, $\mathscr{T} = R_+ = [0,\infty)$, 我们称 $\{X_t\}_{t\in\mathscr{T}}$ 为一个随机过程.

例 1 如果我们取 $\Omega = [0,1]$, $\mathscr{F}$ 为 $[0,1]$ 上的 Lebesgue σ 代数, P 是 $\mathscr{F}$ 上的 Lebesgue 测度, $(\Omega, \mathscr{F}, P)$ 是一个概率空间,我们称为 Lebesgue 概率空间. 作函数:

$$X_t(\omega) = \begin{cases} 0, & 当\ \omega \in [0,\ e^{-t}) \\ 1, & 当\ \omega \in [e^{-t},\ 1] \end{cases}$$

$t \in R_+$, 则 $X_t(\omega)$ 是 Lebesgue 概率空间上的一个随机过程.

这个随机过程的特点，是几乎处处趋于1.

随机过程也可以另一个观点来引进，即从无穷维空间上测度[1]的观点引进。设 $\mathscr{T}$ 为任一实数集，试定义可测空间 $(R_{0,\mathscr{T}}, \mathscr{B}_{\mathscr{T}})$:

令 $\lambda(t)$, $t\in\mathscr{T}$ 是以 $\mathscr{T}$ 为定义域的实值函数，全体这样的函数构成了一个基本空间 $R_{\mathscr{T}}$，即

$$R_{\mathscr{T}}=\{\lambda(t):\ t\in\mathscr{T}\}$$

因此 $R_{\mathscr{T}}$ 中的一个点是 $\mathscr{T}$ 上的一个实值函数。任取 n 个实值 $\lambda_1,\cdots,\lambda_n$，考虑 $R_{\mathscr{T}}$ 的子集:

$$W_n=\{\lambda(t):\lambda(t_1)<\lambda_1,\cdots,\lambda(t_n)\leqslant\lambda_n\}$$

这里 $t_i\in\mathscr{T}$ 任意固定 $(i=1,\ \cdots,\ n)$。可见 W_n 实际上依赖于 $t_1,\cdots,t_n;\ \lambda_1,\cdots,\lambda_n$.现在令 n, t_i, λ_i 分别变动，便得到一族子集类 $W=\{W_n\}$.定义

$$\mathscr{B}_{\mathscr{T}}=\sigma(W)$$

即包含W的最小σ代数。我们称 $\mathscr{B}_{\mathscr{T}}$ 中的任一集为 $R_{\mathscr{T}}$ 中的一个 Borel 可测集。任何一个随机过程 $\{X_t\}_{t\in\mathscr{T}}$，总要在 $(R_{\mathscr{T}}, \mathscr{B}_{\mathscr{T}})$ 上引进一个概率测度：首先定义 W_n 的概率测度为

$$P(W_n)=P(\omega: X_{t_1}\leqslant\lambda_1,\cdots,\ X_{t_n}\leqslant\lambda_{t_n})\qquad(3.1.1)$$

然后将W上的测度扩张到 $\sigma(W)$ 上去。

下面，我们不加证明地介绍随机过程的存在定理。为此我们先引入 n 维分布函数的概念:

n 维分布函数 $F_n(x_1, x_2,\cdots,x_n)$ 是在 n 维欧氏空间 R^n 上定义的实值函数，满足:

(1) 对于任意的区间 $I_k=[a_k,\ b_k)$, $1\leqslant k\leqslant n$

$$\Delta_{I_1}\cdots\Delta_{I_n}F_n(x)\geqslant 0\qquad(3.1.2)$$

---

1) 无穷维空间上的测度在附录A中给出。

其中 $x=(x_1,\cdots,x_n)$;

$$\Delta_{I_k}F(x)=F(x_1,\cdots,x_{k-1},b_k,x_{k+1},\cdots,x_n)-F(x_1,\cdots,x_{k-1},a_k,x_{k+1},\cdots,x_n);$$

(2) 如果当 k 上升地趋于 ∞,记作 $k\uparrow\infty$, $x_j^{(k)}\downarrow x_j(1\leqslant j\leqslant n)$, 则当 $k\to\infty$ 时

$$F_n(x_1^{(k)},\cdots,x_n^{(k)})\to F_n(x_1,\cdots,x_n)$$

(3) 如果对某一 j, $x_j\downarrow-\infty$, 则 $F_n(x_1,\cdots,x_n)\downarrow 0$, 并且如果对所有 $j(1\leqslant j\leqslant n)$, $x_j\uparrow\infty$, 则

$$F_n(x_1,\cdots,x_n)\uparrow 1$$

现在讨论 d 维随机过程,设有一族有穷维分布函数:

$$F=\{F_{t_1,\cdots,t_n}(\lambda_1,\cdots,\lambda_n)\},\ t\in\mathscr{T},\ \lambda_i\in R^d$$

如果它们满足下列两条件,则我们称它们是相容的:

(1) 对于 $(1,\cdots,n)$ 的任一排列 $(\alpha_1,\cdots,\alpha_n)$, 有

$$F_{t_1,\cdots,t_n}(\lambda_1,\cdots,\lambda_n)=F_{t\alpha_1\cdots t\alpha_n}(\lambda_{\alpha_1}\cdots\lambda_{\alpha_n})$$

(2) 如果 $m<n$, 则

$$F_{t_1,\cdots,t_m}(\lambda_1,\cdots,\lambda_m)=\lim_{\lambda_{m+1},\cdots,\lambda_n\to\infty}F_{t_1,\cdots,t_n}(\lambda_1,\cdots,\lambda_n)$$

定理 3.1.1 (随机过程存在定理) 设 F 是满足相容性条件的有穷维分布函数族,那么必存在概率空间 $(\Omega,\mathscr{F},P)$ 及定义于其上的随机过程 $\{X_t(\omega)\}_{t\in\mathscr{T}}$, 使 $\{X_t(\omega)\}_{t\in\mathscr{T}}$ 的有穷维分布函数族与 F 相重合,即

$$P(\omega:X_{t_1}\leqslant\lambda_1,\cdots,X_{t_n}\leqslant\lambda_n)=F_{t_1,\cdots,t_n}(\lambda_1,\cdots,\lambda_n)$$

在随机过程中,除了鞅外,马尔可夫过程和平稳过程也是两类最常见的过程。

定义 3.1.1 设 $\{\xi(t)\}_{t\in\mathscr{T}}$ 是随机过程,如果它的数学期望 $E\xi_t=a$ 不依赖于 t,它的相关阵

$$B(t,s)=E[\xi(t)\xi^T(s)]$$

只依赖于差 $t-s=\tau$, 即 $B(t,s)=B(\tau)$, 则称 $\{\xi(t)\}$

为弱平稳过程。当 $\mathscr{T}$ 为整数集时，称 $\{\xi(t)\}_{t\in\mathscr{T}}$ 为弱平稳序列。

定义 3.1.2 随机过程 $\{\xi(t)\}_{t\in\mathscr{T}}$ 称为强平稳的，如果时间起点的推移不影响它的统计特性，即两个随机过程 $\xi(t)$ 和 $\xi(t+\varepsilon)$，对任给 ε 均具有同样的有穷维分布. 由定义可知，在均方差存在的情况下，强平稳过程一定是弱平稳过程。当 $\mathscr{T}$ 为整数集时，$\{\xi(t)\}$ 称为强平稳序列。

定义 3.1.3 设 $\{X_t(\omega)\}_{t\in\mathscr{T}}$ 为定义在概率空间 $(\Omega,\mathscr{F},P)$上并取值于 R^d 中的随机过程. 如果对任意有限个 $t_1<\cdots<t_n\in\mathscr{T}$ 和任意 $A\in\mathscr{B}(R^d)$ 有

$$P(X_{t_n}\in A\,|\,X_{t_1},\cdots,X_{t_{n-1}})=P(X_{t_n}\in A\,|\,X_{t_{n-1}}),\ \text{a.s.} \tag{3.1.3}$$

则称此过程为马尔可夫过程，性质 (3.1.3) 称为马氏性。

有关马尔可夫过程与平稳随机过程的详细内容，在一般随机过程的书中都能找到[30]，这里只介绍它们的概念。

习　题　3.1

1. 设 ξ 是随机变量，$E\xi^4<C<\infty$，于是

$$P(|\xi|\geqslant\varepsilon)\leqslant\frac{E\xi^4}{\varepsilon^4}$$

3.2 布朗运动与维纳过程

提要 本节引进布朗运动和维纳过程的定义，并给出布朗运动是连续过程的严格证明。

1827 年布朗发现在水中的花粉（或其他液体中的某些微粒）在不停地运动。分子的运动会产生一种涨落不定的

力. 粒子每秒钟所受的碰撞次数达到 10^{21} 次. 因此可以认为粒子是受到了很多微小随机力的作用而产生的随机运动. 这种随机运动称为布朗运动. 后来维纳分析了这一过程, 作了如下的解释:以 X_t 表示粒子在 t 时刻所在位置的一个坐标, 假设液体是均匀的. 此时会自然设想自时刻 t_1 到 t_2 的位移 $X_{t_2}-X_{t_1}$ 是许多几乎独立的小位移之和,即许多小随机变量之和.根据中心极限定理,自然应假设 $X_{t_2}-X_{t_1}$ 有正态分布. 由液体的均匀性, 还会自然地假设 $E(X_{t_2}-X_{t_1})=0$, 而其方差则为 $E(X_{t_2}-X_{t_1})^2=\sigma^2(t_2-t_1)$. 这里 $\sigma>0$ 是依赖于液体具体性质的常数,由液体的均匀性知,它应与时间 t 及空间位置无关. 又因位移 $X_{t_n}-X_{t_{n-1}},\cdots, X_{t_2}-X_{t_1}$ 分别为许多几乎独立的小位移之和,故应设它们互相独立.

把上面的分析写成下面的严格定义:

定义 3.2.1 维纳过程(或称布朗运动)是指满足下列条件的随机过程 $X_t(t\in R_+)$:

(1) $X_0=0$;

(2) X_t 是独立增量过程;

(3) 如果 $0\leqslant s<t$, 则 X_t-X_s 服从正态分布, 并且 $E(X_t-X_s)=0$, 存在常数 $\sigma>0$, $E(X_t-X_s)^2=\sigma^2(t-s)$.

当 $\sigma=1$ 时,$\{X_t\}$ 称为标准布朗运动. 设 X_t 是 ν 维过程,如果各维都是互相独立的标准布朗运动,则称 X_t 为 ν 维标准布朗运动.

容易证明布朗运动是鞅(见习题 3.2.1).

进一步可以证明,布朗运动是一个连续过程. 为此,我们需要引进一个由 Kolmogorov 证明了的引理.

引理 3.2.1 (Kolmogorov 引理) 设 $X_t(t\in R_+)$ 为 ν 维随机过程. 如果有三个正常数 c, α, β 使得

$$E|X_t - X_s|^\beta \leqslant c|t - s|^{1+\alpha} \qquad (3.2.1)$$

则存在 X_t 的一个修正过程(关于修正过程的定义,参见 2.9 节),其轨道连续。

证明 为了简单起见,我们只考虑 $t\in[0,1]$ 的情况. 对于 $t\in R_+$, 可以作同样的证明。

设 $0<\nu<\dfrac{\alpha}{\beta}$, δ 为正数且使

$$(1-\delta)\ (\alpha+1-\beta\nu) > 1+\delta \qquad (3.2.2)$$

这样的 δ 总是可以取到的。

由式 (3.2.1) 知, $P\{\omega|$ 存在某一对 $0\leqslant i\leqslant j\leqslant 2^n$, $j-i\leqslant 2^{n\delta}$ 使得

$$|X_{j2^{-n}} - X_{i2^{-n}}| > [(j-i)2^{-n}]^\nu\}$$

$$\leqslant \sum_{\substack{0\leqslant i\leqslant j\leqslant 2^n\\ j-i\leqslant 2^{n\delta}}} P[|X_{j2^{-n}} - X_{i2^{-n}}| > [(j-i)2^{-n}]^\nu]$$

$$\leqslant \sum_{\substack{0\leqslant i\leqslant j\leqslant 2^n\\ j-i\leqslant 2^{n\delta}}} [(j-i)2^{-n}]^{-\nu\beta} E|X_{j2^{-n}} - X_{i2^{-n}}|^\beta$$

$$\leqslant c_1 \sum_{\substack{0\leqslant i\leqslant j\leqslant 2^n\\ j-i\leqslant 2^{n\delta}}} [(j-i)2^{-n}]^{1+\alpha-\beta\nu}$$

$$\leqslant c_1 \sum_{\substack{0\leqslant i\leqslant j\leqslant 2^n\\ j-i\leqslant 2^{n\delta}}} 2^{-n(1-\delta)(1+\alpha-\beta\nu)}$$

$$\leqslant c_2 2^{n[1+\delta-(1-\delta)(1+\alpha-\beta\nu)]} = c_2 2^{-n\mu}$$

由式 (3.2.2), 这里 $\mu>0$, c_1, c_2 均为常数。由于 $\Sigma 2^{-n\mu}<\infty$, 根据引理 2.1.1, 当 $0\leqslant i\leqslant j\leqslant 2^n, j-i\leqslant 2^{n\delta}$, $n\geqslant m(\omega)$ 时

$$|X_{j2^{-n}} - X_{i2^{-n}}| \leqslant [(j-i)2^{-n}]^\nu, \text{ a.s.} \qquad (3.2.3)$$

这里 $m(\omega)$ 是某一充分大的正整数。

上面证明 $t_2 = j2^{-n}$, $t_1 = i2^{-n}$ 时,得出了

$$|X_{t_2} - X_{t_1}| \leqslant |t_2 - t_1|^{\nu}, \text{ a.s.} \tag{3.2.4}$$

下面证明式(3.2.4)对于(0，1)区间中任意两个二进制有理数都成立。

现在设 t_1，t_2 是在区间(0,1)中使得 $t_1 < t_2$ 的任意两个二进制有理数，$\tau = t_2 - t_1 < 2^{-m(1-\delta)}$，这里 $m = m(\omega)$ 与式(3.2.3)中的相同.选取 $n \geqslant m$ 使得

$$2^{-(n+1)(1-\delta)} \leqslant \tau < 2^{-n(1-\delta)}$$

t_1 和 t_2 可表示为

$$t_1 = i2^{-n} - 2^{-p_1} - \cdots - 2^{-p_k}$$

$$t_2 = j2^{-n} + 2^{-q_1} + \cdots + 2^{-q_l}$$

这里 $n < p_1 < \cdots < p_k$，$n<q_1<\cdots<q_l$，于是 $t_1 \leqslant i2^{-n} \leqslant j2^{-n} \leqslant t_2$，同时由于 $j - i \leqslant \tau 2^n < 2^{-n+n\delta} \cdot 2^n = 2^{n\delta}$。因此由式(3.2.3)，可得

$$|X_{(i2^{-n}-2^{-p_1}-\cdots-2^{-p_r})} - X_{(i2^{-n}-2^{-p_1}-\cdots-2^{-p_{r-1}})}| \leqslant (2^{-p_r})^{\nu}$$

于是

$$|X_{t_1} - X_{i2^{-n}}| \leqslant \Sigma_{r=1}^{k}(2^{-p_r})^{\nu}$$

存在常数 c_3，使

$$|X_{t_1} - X_{i2^{-n}}| \leqslant c_3(2^{-n})^{\nu}$$

类似地

$$|X_{t_2} - X_{j2^{-n}}| \leqslant c_3(2^{-n})^{\nu}$$

最后由式(3.2.3)得

$$|X_{j2^{-n}} - X_{i2^{-n}}| \leqslant [(j-i)2^{-n}]^{\nu} \leqslant \tau^{\nu}$$

对于充分大 n 有

$$\begin{aligned} |X_{t_2} - X_{t_1}| &\leqslant |X_{t_2} - X_{j2^{-n}}| + |X_{j2^{-n}} - X_{i2^{-n}}| \\ &\quad + |X_{i2^{-n}} - X_{t_1}| \\ &\leqslant 2c_3(2^{-n})^{\nu} + \tau^{\nu} \\ &\leqslant c(t_2 - t_1)^{\nu} \end{aligned} \tag{3.2.5}$$

其中 c 为某一常数[当 n 充分大时 $c_3(2^{-n})^{\nu} \xrightarrow[n\to\infty]{} 0$]。

设 Q 是(0,1)区间中所有二进制有理数的序列。由式

(3.2.5) 推得，当 $0 \leqslant t_j - t_i \leqslant 2^{-m(1-\delta)}$，$m = m(\omega)$ 时

$$|X_{t_i} - X_{t_j}| \leqslant c(t_j - t_i)^{\alpha} \qquad (3.2.6)$$

因此当 t 限制于 Q 时，$X_t(\omega)$ 是一致连续的。设 $\tilde{X}_t(\omega)$ 为 $X_t(\omega)$ 到 $0 \leqslant t \leqslant 1$ 上的唯一连续扩张，则过程 $\tilde{X}_t(\omega)$ 是连续的。

下面我们证明 $\tilde{X}_t(\omega)$ 是 $X_t(\omega)$ 的一个修正，从而完成本定理的证明。

设 $t^* \in [0, 1]$，并且 $t_j \in Q$，$t_j \to t^*$(当 $j \to \infty$时)。由于 $\tilde{X}_t(\omega)$ 是连续过程，因此

$$\tilde{X}_{t_j}(\omega) \xrightarrow[j \to \infty]{} \tilde{X}_{t^*}(\omega), \text{ a.s.}$$

由式 (3.2.1) 推出 $X_{t_j}(\omega) \xrightarrow[j \to \infty]{P} X_{t^*}(\omega)$。

由于

$$\tilde{X}_{t_j}(\omega) = X_{t_j}(\omega), \text{ a.s.}$$

我们得到 $\tilde{X}_{t^*}(\omega) = X_{t^*}(\omega)$. ▮

从引理的证明已经清楚地看到，X_t 连续过程是指其轨道连续。更确切地说，是存在 X_t 的一个修正过程，其轨道是连续的。

推论 3.2.1 布朗运动存在连续修正.

证明 根据定义 3.2.1 中的 (3)，$X_t - X_s$ 服从正态分布，由式 (2.1.17) 可得

$$E|X_t - X_s|^4 = 3(t-s)^2 \sigma^4, \forall t > s$$

它满足引理 3.2.1，推论得证. ▮

总之，布朗运动也称维纳过程，它是一个具有正态独立增量的过程，其增量的数学期望为零、方差与时间差成正比，$X_0 = 0$。

维纳过程有两个重要性质：

(1) 存在连续轨道的修正过程；

(2) 维纳过程几乎处处不可导（这一点在3.3节详细论述）。

习 题 3.2

1. 证明布朗运动是鞅.

2. 设 σ 是有界停时，$\{W_t\}$ 是维纳过程，$\{\mathscr{F}_t\}$是 $\{W_t\}$ 的适应，令

$$\tilde{W}_t = W_{t\wedge\sigma},\quad \tilde{\mathscr{F}}_t = \mathscr{F}_{t\wedge\sigma}.$$

证明：$\forall t>s$，有

(1) $(\tilde{W}_t,\tilde{\mathscr{F}}_t)$ 是鞅；

(2) $E[(\tilde{W}_t - \tilde{W}_s)|\tilde{\mathscr{F}}_s) = 0$；

(3) $E[(\tilde{W}_t - \tilde{W}_s)^2|\tilde{\mathscr{F}}_s] = E[(t\wedge\sigma) - (s\wedge\sigma)|\tilde{\mathscr{F}}_s]$.

3.3 阶梯函数的随机积分

提要 本节给出了阶梯过程(函数)的定义，阶梯过程关于维纳过程的随机积分定义及其性质；还定义了适应可测过程和循序可测过程.

随机积分是指形如 $\int_0^T H_s dX_s$ 的积分，这里 H_t 和 X_t 都是随机过程. 1944年伊藤最早定义了适应可测过程关于布朗运动(维纳过程)的积分，即

$$I(T) = \int_0^T H_t(\omega)dW_t$$

其中 W_t 是布朗运动，$H_t(\omega)$ 是随机过程. 在本章里 $I(T)$ 的定义及其性质，是我们研究的主要课题. 这一随机积分的重要特点是，积分所得到的过程为一鞅.

严家安指出："1967年 Kunita 和 Watanabe 定义了适应可测过程对一般平方可积鞅的随机积分，迈出了现代随机积分理论的关键性的一步."1970年 Doleans-Dade 和 Meyer

研究了有界可料过程对局部鞅和半鞅的随机积分. 1976 年 Meyer 研究了可选过程对局部鞅的随机积分. 1979 年 Jacod 讨论了非有界可料过程对半鞅的随机积分.

在本书中，我们主要研究伊藤的随机微积分理论。由于人们已经证明，布朗运动的轨线有

$$P\left\{\limsup_{0\leqslant s-t=h\downarrow 0}\frac{|W_t-W_s|}{\sqrt{2h\ln(1/h)}}=1\right\}=1$$

因此 W_t 以概率 1 处处不可微，在一般 Lebesgue-Stieltjes 意义下，积分 $\int_0^T f(t,\omega)dW_t$ 是不能定义的。于是，我们从最简单的阶梯函数开始考虑。

为了使随机积分具有确定的意义，先讨论下列引理。

引理 3.3.1 设 $0=t_0^{(n)}<t_1^{(n)}<\cdots<t_n^{(n)}=t$ 是区间 $[0,t]$ 的一个分划. $W=(W_t,\mathscr{F}_t)$ 是标准维纳过程，那么当 $n\to\infty$, $\max_i[t_{i+1}^{(n)}-t_i^{(n)}]\to 0$ 时

$$\lim_{n\to\infty}E\left|\sum_{i=0}^{n-1}[W_{t_{i+1}^{(n)}}-W_{t_i^{(n)}}]^2-t\right|^2=0$$

即 $\sum_{i=0}^{n-1}[W_{t_{i+1}^{(n)}}-W_{t_i^{(n)}}]^2$ 在 L_2 意义下收敛于 t，记作

$$\operatorname*{l.i.m}_{n}\sum_{i=0}^{n-1}[W_{t_{i+1}^{(n)}}-W_{t_i^{(n)}}]^2=t \tag{3.3.1}$$

且

$$\lim_{n\to\infty}\sum_{i=0}^{n-1}[W_{t_{i+1}^{(n)}}-W_{t_i^{(n)}}]^2=t,\ \text{a.s.} \tag{3.3.2}$$

证明 因为 $\{W_t\}$ 是标准的维纳过程：

$$E\sum_{i=0}^{n-1}[W_{t_{i+1}^{(n)}}-W_{t_i^{(n)}}]^2=\sum_{i=0}^{n-1}(t_{i+1}^{(n)}-t_i^{(n)})=t \tag{3.3.3}$$

为了证明式 (3.3.1)，我们只要验证下式成立就足够了：

$$\text{当 } n\to\infty \text{时，} \operatorname{Var}\sum_{i=0}^{n-1}[W_{t_{i+1}^{(n)}}-W_{t_i^{(n)}}]^2\to 0$$

验证如下：由于维纳过程的增量的独立高斯(正态分布)性质以及 $\operatorname{Var}[W_t-W_s]=t-s,\ \forall t\geqslant s$，故

$$\operatorname{Var}\sum_{i=0}^{n-1}[W_{t_{i+1}^{(n)}}-W_{t_i^{(n)}}]^2=\sum_{i=0}^{n-1}\operatorname{Var}[W_{t_{i+1}^{(n)}}-W_{t_i^{(n)}}]^2$$

$$=2\sum_{i=0}^{n-1}[t_{i+1}^{(n)}-t_i^{(n)}]^2$$

$$\leqslant 2t\max_i[t_{i+1}^{(n)}-t_i^{(n)}]\xrightarrow[n\to\infty]{}0$$

所以式 (3.3.1) 成立。

往证式 (3.3.2)：为了简单起见，假定 $t_i^{(n)}=it/n$（一般情形写起来繁复些，其理由是一样的）. 证明式 (3.3.2) 需要用到下面的引理。

引理 3.3.2 设 $\{\xi_n, n=1,2,\cdots\}$ 是随机变量序列，对任给 $\varepsilon>0$ 有

$$\sum_{n=1}^{\infty}P\{|\xi_n|>\varepsilon\}<\infty \tag{3.3.4}$$

那么当 $n\to\infty$ 时

$$\xi_n\to 0,\ \text{a.s.}$$

证明 设 $A_n^\varepsilon=\{\omega: |\xi_n|>\varepsilon\}$， $B^\varepsilon=\limsup\limits_n A_n^\varepsilon=\bigcap\limits_{n=1}^{\infty}\bigcup\limits_{m=n}^{\infty}A_m^\varepsilon$. 根据引理 2.1.1，若

$$A^*=\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty}A_m$$

$$\sum_{i=0}^{\infty} P(A_i) < \infty \tag{2.1.2}$$

则

$$P(A^*) = 0 \tag{2.1.3}$$

即

$$P(B^{\varepsilon}) = 0$$

又因为

$$\{\omega : \xi_n \nrightarrow 0\} = \bigcup_k B^{1/k}$$

所以

$$P\{\omega : \xi_n \underset{n\to\infty}{\nrightarrow} 0\} = 0$$

即

$$\xi_n \xrightarrow[n\to\infty]{} 0, \text{ a.s.}$$ ∎

现在再来证明式(3.3.2)：因为 $t_i^{(n)} = it/n$，令

$$\xi_n = \sum_{i=0}^{n-1} \left\{ [W_{\frac{(i+1)t}{n}} - W_{\frac{it}{n}}]^2 - \frac{t}{n} \right\} \tag{3.3.5}$$

根据马尔可夫不等式 $\left(E|\xi_n|^4 \geqslant \int_{|\xi_n|>\varepsilon} |\xi_n|^4 dP \geqslant \varepsilon^4 P(|\xi_n| > \varepsilon)\right)$，有

$$P(|\xi_n| > \varepsilon) \leqslant \frac{E|\xi_n|^4}{\varepsilon^4}$$

利用维纳过程在不重叠区间上增量的独立性和正态分布随机变量偶数阶中心距公式，于是有

$$E\xi_n^2 = E\left\{ \sum_{i=0}^{n-1} [W_{\frac{(i+1)t}{n}} - W_{\frac{it}{n}}]^2 \right\}^2$$

$$- 2tE \sum_{i=0}^{n-1} [W_{\frac{(i+1)t}{n}} - W_{\frac{it}{n}}]^2 + t^2$$

$$= 3\frac{t^2}{n} + t^2 - \frac{t^2}{n} - 2t^2 + t^2 = \frac{2t^2}{n}$$

$$= 2\left(\frac{t}{n}\right)t$$

同理有

$$E\xi_n^4 \leqslant C\left(\frac{t}{n}\right)^3 t$$

所以级数

$$\sum_{n=1}^{\infty} P(|\xi_n| > \varepsilon) < \infty$$

根据引理 3.3.2，$\xi_n \xrightarrow[n\to\infty]{} 0$。在假设 $t_j^{(n)} = \frac{jt}{n}$ 下就证明了式(3.3.2)。

注记 我们把方程(3.3.1)和(3.3.2)象征性地记作

$$\int_0^t (dW_s)^2 = \int_0^t ds \qquad (3.3.6)$$

为了定义随机积分 $\int_0^t H_s(\omega)dW_s$，我们需对随机函数类作一些规定，并从最简单函数的随机积分讨论起。

定义 3.3.1 设 $\{H_t(\omega)\}_{t\in R_+}$ 是一个随机过程（随机函数），如果把 $H_t(\omega)$ 看作 (t,ω) 的二元函数，$H_t(\omega)$ 关于乘积测度空间 $(R_+\times\Omega, \mathscr{B}(R_+)\times\mathscr{F}, d\lambda\times dP)$ 可测，则我们称 $H_t(\omega)$ 是一个可测过程或可测随机函数。这里 $\mathscr{B}(R_+)$ 是 R_+ 上的 Borel 代数，λ 是其上的 Lebesgue 测度。

定义 3.3.2 如果对于每一个 t，$H_t(\omega)$ 关于 $\mathscr{F}_t$ 可测，则称 $H_t(\omega)$ 关于 $\{\mathscr{F}_t\}$ 是适应的，$H_t(\omega)$ 称为适应过程或适应函数。

下面引进阶梯过程的概念：

定义 3.3.3 有限区间 $[0, T]$ 存在一个分划 $0 = t_0 < t_1 < \cdots < t_n = T$；$\alpha, \alpha_1, \cdots, \alpha_n$ 分别为 $\mathscr{F}_{t_0}, \mathscr{F}_{t_1}, \cdots, \mathscr{F}_{t_n}$ 可测的随机变量。α 是 $\mathscr{F}_0$ 可测的，若

$$e(t, \omega) = \alpha \chi_{\{0\}}(t) + \sum_{i=0}^{n-1} \alpha_i \chi_{(t_i, t_{i+1}]}(t) \tag{3.3.7}$$

则称 $e(t,\omega)$ 是阶梯函数或阶梯过程

显然阶梯函数 $e(t)$ 关于 $F = \{\mathscr{F}_t\}_{0 \leqslant t \leqslant T}$ 是适应的。

如果 $a(t)$ 是非降右连续函数（如定分布函数），那么

$$\int_0^{\infty} \chi_{[\mu, \nu]} da(t) = a(\nu) - a(\mu) \tag{3.3.8}$$

定义 3.3.4 下式是阶梯函数 (3.3.7) 关于维纳过程的积分：

$$I_t(e) = \alpha W_0 + \sum_{\{0 \leqslant i \leqslant m : t_{m+1} < t\}} \alpha_i [W_{t_{i+1}} - W_{t_i}] + \alpha_{m+1}[W_t - W_{t_{m+1}}]$$

因为 W_t 是维纳过程

$$P(W_0 = 0) = 1$$

上式可写成

$$I_t(e) = \sum_{\{0 \leqslant i \leqslant m : t_{m+1} < t\}} \alpha_i [W_{t_{i+1}} - W_{t_i}] + \alpha_{m+1}[W_t - W_{t_{m+1}}] \tag{3.3.9}$$

为了记法简单，上式右边的和式可写成积分形式，即

$$I_t(e) = \int_0^t e(s, \omega) dW_s \tag{3.3.10}$$

积分 $\int_s^t e(u, \omega) dW_u$ 可以理解为 $I_t(\tilde{e})$，即式 (3.3.10) 右端被积函数 $e(u,\omega)$ 用 $\tilde{e}(u,\omega)$ 代替，其中

$$\tilde{e}(u, \omega) = e(u, \omega) \chi_{\{u > s\}}$$

从阶梯函数随机积分的定义(3.3.9)和(3.3.10),可以推导出随机积分的下列性质:

(1) 设 a, b 是两个常数,则

$$I_t(ae_1 + be_2) = aI_t(e_1) + bI_t(e_2)$$

(2) 若 $0 \leqslant u < t$, 那么

$$\int_0^t e(s,\omega)dW_s = \int_0^u e(s,\omega)dW_s + \int_u^t e(s,\omega)dW_s$$

(3) $I_t(e)$ 是 t 的连续函数 $(0 \leqslant t \leqslant T)$.

这一点可以由 W_t 轨线连续性推得(习题 3.3,1)

(4) $I_t(e)$ 是鞅,即 $\forall t > s$

$$E\left(\int_0^t e(u,\omega)dW_u \,|\, \mathscr{F}_s\right) = \int_0^s e(u,\omega)dW_u, \quad \text{a.s.}$$

证明 由性质(2)知

$$\begin{aligned}
E\left(\int_0^t e(u,\omega)dW_u \,|\, \mathscr{F}_s\right)
&= E\left(\int_0^s e(u,\omega)dW_u \,|\, \mathscr{F}_s\right) \\
&\quad + E\left(\int_s^t e(u,\omega)dW_u \,|\, \mathscr{F}_s\right)
\end{aligned}$$

由于 $\int_0^s e(u,\omega)dW_u$ 是关于 $\mathscr{F}_s$ 可测的随机变量之和,所以

$$\int_0^s e(u,\omega)dW_u \text{ 关于 } \mathscr{F}_s \text{ 可测}$$

又因为

$$\begin{aligned}
E\left(\int_s^t e(u,\omega)dW_u \,|\, \mathscr{F}_s\right) &= E(\alpha_0[W_{t_1} - W_s] \,|\, \mathscr{F}_s) \\
&\quad + E\left(\sum_{1 \leqslant i \leqslant m, t_1 > s, t_m < t} E(\alpha_i[W_{t_{i+1}} - W_{t_i}] \,|\, \mathscr{F}_{t_i}) \,|\, \mathscr{F}_s\right) \\
&\quad + E\{E(\alpha_m[W_t - W_{t_m}] \,|\, \mathscr{F}_{t_m}) \,|\, \mathscr{F}_s\} \\
&= 0
\end{aligned}$$

由此可得

$$E\left(\int_0^t e(u,\omega)dW_u \mid \mathscr{F}_s\right)=\int_0^s e(u,\omega)dW_u, \text{ a.s.}$$ ▌

(5) $$E\left(\int_0^t e_1(u,\omega)dW_u\right)\left(\int_0^t e_2(u,\omega)dW_u\right)$$

$$=E\int_0^t e_1(u,\omega)e_2(u,\omega)du$$

证明留作习题。

(6) 如果对所有 s，$0\leqslant s\leqslant T$ 和 $\omega\in A\in\mathscr{F}_T$

$$e(s,\omega)=0$$

那么

$$\int_0^t e(s,\omega)dW_s=0,\ \forall t\leqslant T,\ \omega\in A$$

定义 3.3.5 $X=(X_t,\ \mathscr{F}_t)$ 称为循序可测的，若对任何 $t\in T$

$$\{(\omega,s):s\leqslant t,\ X_s(\omega)\in B\}\in\mathscr{F}_t\times\mathscr{B}[0,t]$$

这里 $B\in\mathscr{B}(R_1)$。

(7) $I_t(e)=\int_0^t e(u,\omega)dW_u$ 循序可测且有

$$E\int_0^t e(u,\omega)dW_u=0$$

证明留作习题。

习　题　3.3

1. 证明 $I_t(e)$ 是 t 的连续函数。

2. 证明性质 (5)。

[提示：利用 $\alpha_i\in\mathscr{F}_{t_i}$ 和 $E(E(\xi|\mathscr{F}))=E\xi$ 证明]

3. 证明 $I_t(e)$ 循序可测且 $EI_t(e)=0$。

[提示：利用

$$E\alpha_i[W_{t_{i+1}}-W_{t_i}]=E\{\alpha_i E[(W_{t_{i+1}}-W_{t_i})|\mathscr{F}_{t_i}]\}$$

证明]

3.4 有界循序可测函数的随机积分

提要 本节讨论有界循序可测过程关于维纳过程的积分以及这一随机积分的基本性质.在本节中,我们首先证明一个有界循序可测过程可以用一平方可积的适应阶梯过程来逼近。然后再证明平方可积适应阶梯过程列的随机积分以鞅为极限。这样,就把这一鞅过程定义为有界循序可测过程关于维纳过程的积分。

下面我们讨论可以用阶梯过程逼近的随机过程(随机函数)。

引理 3.4.1 设 $H_t(\omega)$ 是一个有界连续适应过程。给定 $t\in R_+$,设 $0=t_0^{(m)}<t_1^{(m)}<\cdots<t_{N(m)}^{(m)}=t$ 是 $[0,t]$ 的一列分划,$\max|t_{n+1}^{(m)}-t_n^{(m)}|\xrightarrow[m\to\infty]{}0$,则

$$E\int_0^t|H_s-H_s^{(m)}|^2ds\xrightarrow[m\to\infty]{}0$$

此处 $H_s^{(m)}=H_{t_n}$(当 $t_n\leqslant s<t_{n+1}$ 时),$H_t^{(m)}=H_{t_{(N(m)-1)}}$.

证明 对固定的 ω,$H_s(\omega)$,$\forall s\in[0,T]$ 是 s 的连续函数。对任给 $\varepsilon>0$,$s\in[t_i,t_{i+1})$,只要 $|t_{i+1}-t_i|<\delta$,$\delta>0$ 必有

$$|H_s^{(m)}-H_s(\omega)|<\varepsilon$$

因此对几乎所有固定的 ω,$H_s^{(m)}(\omega)\xrightarrow[m\to\infty]{}H_s(\omega)$,$\forall s\in[0,t]$,并且 $H_s^{(m)}$ 和 H_s 一致有界。这样,用控制收敛定理,本引理得证。

注意:这里 $\{H_s^{(m)}\}$ 是阶梯函数列。

引理 3.4.2 设 $F_t(\omega)$ 是有界循序可测过程,定义

$$F_s^{(m)} = m \int_{s-\frac{1}{m}}^{s} F_u(\omega) du \tag{3.4.1'}$$

此处当 $u < 0$ 时，令 $F_u(\omega) = 0$，于是 $F_s^{(m)}$ 是有界连续适应过程，并且对每个固定的 ω，几乎所有的 s，$F_s^{(m)}(\omega) \xrightarrow[m\to\infty]{} F_s(\omega)$。进一步，对固定的 t，有

$$E \int_0^t |F_s - F_s^{(m)}|^2 ds \xrightarrow[m\to\infty]{} 0$$

证明 $F_s^{(m)}$的连续性出自 Lebesgue 不定积分的连续性，$F_s^{(m)} \in \mathscr{F}_s$ 来自 Fubini 定理。由突变函数论可知 $F_s^{(m)} \xrightarrow[m\to\infty]{} F_s$。应用控制收敛定理得最后的均方收敛性。 ▌

定理 3.4.1 若 F_s 是一个有界的循序可测过程，则存在一列平方可积的适应阶梯过程 $H_s^{(m)}$，使对每个固定的 t，有

$$E \int_0^t |H_s^{(m)} - F_s|^2 ds \xrightarrow[m\to\infty]{} 0$$

证明 先按引理 3.4.2 作出 $F_s^{(m)}$ 且使

$$E \int_0^t |F_s^{(m)} - F_s|^2 ds \xrightarrow[m\to\infty]{} 0$$

假定(否则可挑选子序列)

$$E \int_0^t |F_s^{(m)} - F_s|^2 ds < \frac{1}{m}$$

再按引理 3.4.1，对每个 $F_s^{(m)}$ 作出子序列 $H_s^{(m,n)}$，使 $H_s^{(m,n)}$ 成为平方可积的适应阶梯过程，并且

$$E \int_0^t |H_s^{(m,n)} - F_s^{(m)}|^2 ds < \frac{1}{n}$$

这样，我们得到

$$\begin{array}{llll}
H_s^{(1,1)}, & H_s^{(1,2)}, & H_s^{(1,3)}, & \cdots \\
H_s^{(2,1)}, & H_s^{(2,2)}, & H_s^{(2,3)}, & \cdots \\
H_s^{(3,1)}, & H_s^{(3,2)}, & H_s^{(3,3)}, & \cdots \\
\cdots & \cdots & &
\end{array}$$

我们记 $H_t^{(m)} = H_t^{(m,m)}$，则 $H_t^{(m)}$ 为平方可积适应阶梯过程，并且

$$E\int_0^t |H_s^{(m)} - F_s|^2 ds \leqslant 2E\int_0^t |H_s^{(m,m)} - F_s^{(m)}|^2 ds + 2E\int_0^t |F_s^{(m)} - F_s|^2 ds \leqslant 2\frac{1}{m} + 2\frac{1}{m} = \frac{4}{m} \xrightarrow[m\to\infty]{} 0 \qquad \blacksquare$$

引理 3.4.3 设 $H_t^{(m)}$ 是一列平方可积的适应阶梯过程，$E\int_0^t |H_s^{(m)} - H_s^{(n)}|^2 ds \xrightarrow[m,n\to\infty]{} 0$，$\forall t \in R_+$，那么存在一个鞅 M_t，$E|M_t|^2 < \infty$，$\forall t \in R_+$ 且

$$E\left|M_t - \int_0^t H_s^{(m)} dW_s\right|^2 \xrightarrow[m\to\infty]{} 0 \qquad (3.4.2)$$

证明 利用阶梯函数随机积分的性质 (5)

$$E\left|\int_0^t H_s^{(m)} dW_s - \int_0^t H_s^{(n)} dW_s\right|^2 = E\left|\int_0^t (H_s^{(m)} - H_s^{(n)}) dW_s\right|^2 = E\int_0^t |H_s^{(m)} - H_s^{(n)}|^2 ds \xrightarrow[m,n\to\infty]{} 0$$

于是 $\int_0^t H_s^{(m)} dW_s$ 构成 $L_2(\Omega, \mathscr{F}, P)$ 中的 Canchy 序列。由于 L_2 空间的完备性，可知存在 $M_t \in L_2$，使式 (3.4.2) 成立。

往证，M_t 是 $\{\mathscr{F}_t\}$ 鞅．实际上对 $t > s$，和每个 $A \in \mathscr{F}_s$，

$$\int_A M_t dP = \lim_{m\to\infty} \int_A \left(\int_0^t H_u^{(m)} dW_u\right) dP = \lim_{m\to\infty} \int_A \left(\int_0^s H_u^{(m)} dW_u\right) dP \ [\text{性质}(4)] = \int_A M_s dP$$

故 $(M_t, \mathscr{F}_t)$ 为鞅。 ∎

定义 3.4.1 设 $F_t(\omega)$ 是一个有界的循序可测过程，令 $H_t^{(m)}$ 为任何满足定理 3.4.1 的阶梯过程。由引理 3.4.3 知，$\int_0^t H_s^{(m)} dW_s$，当 $m \to \infty$ 时存在均方极限 M_t，我们定义

$$\int_0^t F_s dW_s = M_t$$

为 F_t 关于维纳过程的随机积分。

很明显，上述定义与所选择的逼近序列 $H_t^{(m)}$ 无关。事实上，设 $H_t^{(m)}$ 和 $\overline{H}_t^{(m)}$ 是两个阶梯过程，并使

$$E\int_0^t |F_s - H_s^{(m)}|^2 ds \xrightarrow[m\to\infty]{} 0$$

$$E\int_0^t |F_s - \overline{H}_s^{(m)}|^2 ds \xrightarrow[m\to\infty]{} 0$$

于是

$$E\int_0^t \left| H_s^{(m)} - \overline{H}_s^{(m)} \right|^2 ds \to 0$$

这样，即得到

$$\begin{aligned} E\left|\int_0^t H_s^{(m)} dW_s - \int_0^t \overline{H}_s^{(m)} dW_s\right|^2 &= E\left|\int_0^t (H_s^{(m)} - \overline{H}_s^{(m)}) dW_s\right|^2 \\ &= E\int_0^t (H_s^{(m)} - \overline{H}_s^{(m)})^2 ds \xrightarrow[m\to\infty]{} 0 \end{aligned}$$

故 $\int_0^t H_s^{(m)} dW_s$ 与 $\int_0^t \overline{H}_s^{(m)} dW_s$ 趋于同一极限。

下面讨论随机积分的基本性质。

定理 3.4.2 设 $\xi_t^{(1)}$，$\xi_t^{(2)}$ 为两个有界循序可测过程，λ_1 和 λ_2 为两个实数，于是

$$\int_0^t [\lambda_1 \xi_s^{(1)} + \lambda_2 \xi_s^{(2)}] dW_s$$

$$= \lambda_1 \int_0^t \xi_s^{(1)} dW_s + \lambda_2 \int_0^t \xi_s^{(2)} dW_s \tag{3.4.3}$$

证明 我们分别用平方可积适应阶梯过程列 $\xi_s^{(1,m)}$ 逼近 $\xi_s^{(1)}$; $\xi_s^{(2,m)}$ 逼近 $\xi_s^{(2)}$. 此时 $\lambda_1\xi_s^{(1,m)} + \lambda_2\xi_s^{(2,m)}$ 仍是适应阶梯过程，而且

$$E \int_0^t |(\lambda_1\xi_s^{(1)} + \lambda_2\xi_s^{(2)}) - (\lambda_1\xi_s^{(1,m)} + \lambda_2\xi_s^{(2,m)})|^2 ds \to 0$$

但从阶梯函数随机积分的性质 (1) 知

$$\int_0^t (\lambda_1\xi_s^{(1,m)} + \lambda_2\xi_s^{(2,m)}) dW_s$$

$$= \lambda_1 \int_0^t \xi_s^{(1,m)} dW_s + \lambda_2 \int_0^t \xi_s^{(2,m)} dW_s$$

两边同时使 $m \to \infty$，定理即得证。 ∎

定理 3.4.3 设 ξ_s 是任意有界循序可测过程，于是

$$E\left(\int_0^t \xi_s dW_s\right)^2 = E \int_0^t \xi_s^2 ds \tag{3.4.4}$$

证明 取阶梯过程 $\xi_s^{(m)}$，使 $E \int_0^t |\xi_s^{(m)} - \xi_s|^2 ds \to 0$. 因为

$$\sqrt{E \int_0^t |\xi_s|^2 ds} \leqslant \sqrt{E \int_0^t |\xi_s - \xi_s^{(m)}|^2 ds} + \sqrt{E \int_0^t |\xi_s^{(m)}|^2 ds}$$

所以

$$\sqrt{E \int_0^t |\xi_s|^2 ds} \leqslant \liminf_m \sqrt{E \int_0^t |\xi_s^{(m)}|^2 ds} \tag{3.4.5}$$

又因为

$$\sqrt{E \int_0^t |\xi_s^{(m)}|^2 ds} \leqslant \sqrt{E \int_0^t |\xi_s^{(m)} - \xi_s|^2 ds} + \sqrt{E \int_0^t |\xi_s|^2 ds}$$

故

$$\limsup_{n} \sqrt{E\int_0^t |\xi_s^{(m)}|^2 ds} \leqslant \sqrt{E\int_0^t |\xi_s|^2 ds} \qquad (3.4.6)$$

由式 (3.4.5) 和 (3.4.6) 得

$$E\int_0^t |\xi_s|^2 ds = \lim_m E\int_0^t |\xi_s^{(m)}|^2 ds$$

另一方面，根据有界循序可测过程的随机积分的定义，有

$$E\left|\int_0^t \xi_s^{(m)} dW_s - \int_0^t \xi_s dW_s\right|^2 \xrightarrow[m\to\infty]{} 0$$

用上面类似的推导可得

$$E\left|\int_0^t \xi_s dW_s\right|^2 = \lim_{m\to\infty} E\left|\int_0^t \xi_s^{(m)} dW_s\right|^2$$

再根据阶梯函数随机积分的性质 (5)

$$E\left|\int_0^t \xi_s^{(m)} dW_s\right|^2 = E\int_0^t |\xi_s^{(m)}|^2 ds$$

于是式 (3.4.4) 得证。 ∎

定理 3.4.4 设 ξ_t 为任何有界循序可测过程，那么

(1) $E\int_0^t \xi_s dW_s = 0$; (3.4.7)

(2) $\int_0^t \xi_s dW_s$ 存在一个连续修正。

证明 由于对阶梯函数 $\xi_t^{(m)}$ 有

$$E\int_0^t \xi_s^{(m)} dW_s = 0$$

故可推得式 (3.4.7) 成立(习题 3.4, 1)。

因为式 (3.4.7) 成立，由定理 2.11.5 知 $\int_0^t \xi_s dW_s$ 的连续修正存在。 ∎

定理 3.4.5 设 Y_t 是 $\int_0^t \xi_s dW_s$ 的连续修正，Y_t 为连续鞅。

证明 取阶梯过程 $\xi_t^{(m)}$ 使 $E\left[\int_0^t |\xi_s^{(m)}-\xi_s|^2 ds\right]\xrightarrow[m\to\infty]{}0$. 根据阶梯函数随机积分的定义,可知 $\int_0^t \xi_s^{(m)}dW_s$ 为连续鞅,并且

$$E\left[\left|Y_t-\int_0^t \xi_s^{(m)}dW_s\right|^2\right]\xrightarrow[m\to\infty]{}0$$

由定理 2.11.5,注意到 $\left|Y_t-\int_0^t \xi_s^{(m)}dW_s\right|$ 对固定的 t,有

$$\begin{aligned}
&E^{\frac{1}{2}}\left[\sup_{0\leqslant s\leqslant t}\left(\left|Y_s-\int_0^s \xi_u^{(m)}dW_u\right|\right)^2\right]\\
&\qquad\leqslant 2\sup_{0\leqslant s\leqslant t}E^{\frac{1}{2}}\left[\left|Y_s-\int_0^s \xi_u^{(m)}dW_u\right|\right]
\end{aligned}$$

后者趋于0,于是可以选择 $\{m\}$ 的子序列 $\{m_k\}_k$,使对几乎所有 ω,有

$$\sup_{0\leqslant s\leqslant t}\left(\left|Y_s-\int_0^s \xi_u^{(m_k)}dW_u\right|\right)\xrightarrow[k\to\infty]{}0$$

这就是说,对几乎所有固定的 ω, $\int_0^s \xi_u^{(m_k)}dW_u(\omega)$ 一致地在 $[0, t]$ 中收敛于 $Y_s(\omega)$. 这样,从 $\int_0^s \xi_u^{(m_k)}dW_u$ 的连续性,就推得了 $Y_s(\omega)$ 的连续性. ∎

以后,我们总用 $\int_0^t \xi_s dW_s$ 表示它的连续修正。这样,随机积分的值的过程总是连续的。

习 题 3.4

1. 证明 $E\int_0^t \xi_s dW_s=0$,其中 ξ_s 是有界循序可测过程.
2. 证明阶梯函数的随机积量是连续过程.
3. 在定理 3.4.1 中,为什么对每一个 $F_t^{(m)}$ 可作出一子序列 $H_t^{(m,n)}$,

$H_t^{(m,n)}$ 平方可积且

$$E\int_0^t |H_s^{(m,n)} - F_s^{(m)}|^2 ds < \frac{1}{n}$$

3.5 局部有界循序可测过程的积分

提要 本节首先讨论随机积分与停时的关系，证明T是有界停时，$H_t(\omega)$是有界循序可测过程时$\chi_{[0,T]}\cdot H_t(\omega)$也是有界循序可测过程。然后定义局部有界的循序可测过程及其随机积分。最后给出一个随机积分的计算实例。

随机积分与停时的关系可以用下列定理来说明。

定理 3.5.1 设T是有界停时，$H_t(\omega)$是有界循序可测过程，那么 $Z_t = \chi_{[0,T]}H_t$ 也是有界循序可测过程，并且

$$\int_0^t Z_s dW_s = \int_0^{t\wedge T} H_s dW_s \tag{3.5.1}$$

证明 因为 $[0,T] = \{(t,\omega): 0 \leqslant t \leqslant T(\omega)\} \in \mathscr{B}[0,t] \times \mathscr{F}_t$，所以 $\chi_{[0,T]}$ 是循序可测过程，而且 Z_t 又是两个有界循序可测过程之积，Z_t 本身也是一个有界的循序可测过程。

我们先构造一列停时 $T_m \downarrow T$，并且每个 T_m 只取有限个值。设 $T \leqslant n$，对每个m定义

$$T_m(\omega) = \sum_{k=0}^{n2^m} (k+1)2^{-m}\chi_{\left[\omega:\frac{k}{2^m}\leqslant T(\omega)<\frac{k+1}{2^m}\right]}(\omega) \tag{3.5.2}$$

我们可以验证，对任何 $t \in R_+$，$[T_m \leqslant t] \in \mathscr{F}_t$，故 T_m 为停时，且 $T_m \downarrow T$。现在定义

$$Z_s^{(m)} = \chi_{[0,T_m]}H_s \tag{3.5.3}$$

如果我们能证明

$$\int_0^t Z_s^{(m)} dW_s = \int_0^{t\wedge T_m} H_s dW_s \tag{3.5.4}$$

那么定理就被证明了. 事实上, 从随机积分的连续性知, 当 $m \to \infty$ 时,右边趋于 $\int_0^{t\wedge T} H_s dW_s$. 再看左边,不失一般性,有

$$E\left[\left|\int_0^t Z_s dW_s - \int_0^t Z_s^{(m)} dW_s\right|^2\right]$$
$$\leqslant E\left[\int_0^t |Z_s - Z_s^{(m)}|^2 ds\right] \leqslant \frac{1}{2^m} \xrightarrow[m\to\infty]{} 0$$

因此当 $m \to \infty$ 时,式 (3.5.4) 左边趋于 $\int_0^t Z_s dW_s$. 这样在式 (3.5.4)的两边取极限,即得定理.

往证式 (3.5.4): 我们只需先对 H_t 为阶梯过程时证明该式, 然后取极限就得到一般情况下的结论. 当 H_t 为阶梯过程时, $Z_t^{(m)}$ 也是阶梯过程. 这样,我们可以对阶梯过程直接验证式 (3.5.4). 我们验证下列事实: 设 Z_t 和 H_t 是两个有界阶梯过程, $\pi=\{0=t_0<t_1<\cdots<t_k=n\}$ 是区间 $[0, n]$ 的一个有限分划. $S(\omega)$ 取值于集合 π, 并且 S 是一个停时. 假定当 $t \leqslant S(\omega)$ 时, $Z_t(\omega)=H_t(\omega)$; 当 $t>S(\omega)$ 时, $Z_t(\omega)=0$, 那么

$$\int_0^t Z_s dW_s = \int_0^{t\wedge S} H_s dW_s \tag{3.5.5}$$

事实上,从式 (3.3.9) 可以清楚地看出,当 $t \leqslant S(\omega)$ 时,两者相同,当 $t > S(\omega)$ 时,两者都等于

$$\sum_{t_j<S(\omega)} H_{t_j}(W_{t_{j+1}} - W_{t_j})$$

故式 (3.5.5) 为真,从而证明了本定理. ∎

定义 3.5.1 设 H_t 是一个循序可测过程, 如果存在有界停时 $T_n \uparrow \infty$, a.s., 使 $\chi_{[0,T_n]}H_t$ 为有界循序可测过程,则称 H_t 为局部有界的循序可测过程.

例如，维纳过程就是有界的循序可测过程。

定义 3.5.2 设 H_t 为局部有界的循序可测过程，那么存在一个连续过程 Y_t，使

$$Y_{t\wedge T_n}=\int_0^t \chi_{[0,T_n]}H_s dW_s, \quad \text{a.s.}$$

Y_t 称为 H_t 关于 W_t 的随机积分，记为

$$Y_t=\int_0^t H_s dW_s$$

由于 $\chi_{[0,T_n]}H_t$ 的有界性

$$\int_0^t \chi_{[0,T_n]}H_s dW_s$$

对每个固定的 n 有定义，而且当 $m>n$ 时，如果 $t<T_n$，则

$$\begin{aligned}
Y_t=Y_{t\wedge T_n}&=\int_0^t \chi_{[0,T_n]}H_s dW_s\\
&=\int_0^t \chi_{[0,T_n]}\chi_{[0,T_m]}H_s dW_s\\
&=\int_0^{t\wedge T_n} \chi_{[0,T_m]}H_s dW_s\\
&=\int_0^t \chi_{[0,T_m]}H_s dW_s=Y_{t\wedge T_m}
\end{aligned}$$

所以上述定义有意义。

另外为使定义含义明确，我们还须证明以下事实：

定理 3.5.2 设 $\{T_n\}$，$\{T'_n\}$ 是两列满足定义 3.5.1 的停时序列，于是它们按定义 3.5.2 所作的随机积分几乎处处相等。

这样，随机积分的定义与 $\{T_n\}$ 选取哪一个无关，只要它满足定义 3.5.2 即可。

证明 事实上，记 $S_n=T_n\wedge T'_n$，于是 $S_n\uparrow\infty$, a.s.，而且 $\chi_{[0,S_n]}H_t$ 仍是有界的。如果记 Z_t 为关于 S_n 按定义 3.5.2 作出的随机积分，Y_t 和 Y'_t 分别为 T_n 和 T'_n 按定义 3.5.2 作出的随机积分，于是由定理 3.5.1，有

$$Y_{t\wedge S_n} = Y_{t\wedge S_n\wedge T_n} = \int_0^t \chi_{[0,S_n]}\chi_{[0,T_n]}H_s dW_s$$

$$= \int_0^t \chi_{[0,S_n]}H_s dW_s = Z_{t\wedge S_n}$$

同理可证 $Y'_{t\wedge S_n} = Z_{t\wedge S_n}$，故 $Y_{t\wedge S_n} = Y'_{t\wedge S_n}$，a.s.，使 $n\to\infty$，$Y_t = Y'_t$，a.s.。 ∎

定义 3.5.3 设 X_t 是一个连续过程，如果存在有界停时列 $T_n \uparrow \infty$，a.s. 使 $X_t^{T_n}(\omega) \triangleq X_{t\wedge T_n}(\omega)$ 为连续鞅，则称 X_t 是一个局部连续鞅。

定理 3.5.3 设 X_t 是按定义 3.5.2 作出的随机积分，于是 X_t 为局部连续鞅。

证明 因为

$$X_t^{T_n} = \int_0^t \chi_{[0,T_n]}H_s dW_s$$

为连续鞅。 ∎

例 1 计算 $\int_a^b W_t dW_t$，这里 W_t 是标准维纳过程.

解 W_t 是局部有界循序可测过程，故随机积分存在. 设 $a = t_0 < t_1 < \cdots < t_n = b$ 是区间 $[a, b]$ 的一个分划，则 $\int_a^b W_t dW_t$ 可看作下列和式的 L_2 极限，且当 $n\to\infty$，$\max_i |t_{i+1}-t_i| \to 0$ 时

$$\sum_{i=0}^{n-1} W_{t_i}[W_{t_{i+1}} - W_{t_i}] \xrightarrow{L_2} \int_a^b W_t dW_t$$

因为

$$\sum_{i=0}^{n-1} W_{t_i}[W_{t_{i+1}} - W_{t_i}] = \frac{W_b^2}{2} - \frac{W_a^2}{2}$$

$$- \frac{1}{2}\sum_{i=0}^{n-1}[W_{t_{i+1}} - W_{t_i}]^2$$

根据引理 3.3.1，当 $\max\limits_i |t_{i+1}-t_i|\to 0$ 时

$$\lim_n \sum_{i=0}^{n-1}[W_{t_{i+1}}-W_{t_i}]^2=b-a$$

故

$$\int_a^b W_t dW_t=\frac{W_b^2}{2}-\frac{W_a^2}{2}-\frac{1}{2}(b-a)$$

习　题　3.5

1.设 $H_t^{(1)}$ 和 $H_t^{(2)}$ 为两个局部有界循序可测过程，λ_1 和 λ_2 为两个实数，证明

$$\int_0^t[\lambda_1 H_s^{(1)}+\lambda_2 H_s^{(2)}]dW_s=\lambda_1\int_0^t H_s^{(1)}dW_s+\lambda_2\int_0^t H_s^{(2)}dW_s$$

2.设

$$H_t(\omega)=\begin{cases}1, & 当\ t\leqslant 1\\ 0, & 当\ t>1\end{cases}$$

写出积分 $\int_0^t H_s dW_s$。

3.求证 $E\left|\int_0^t s\,dW_s\right|\leqslant 3^{-\frac{1}{2}}t^{\frac{3}{2}}$.

$\left[提示:\ E\left|\int_0^t s\,dW_s\right|\leqslant E^{\frac{1}{2}}\left[\left|\int_0^t s\,dW_s\right|^2\right]\right]$

3.6　伊藤微分公式

提要·伊藤微分公式对于随机积分的计算是十分有用的。本节从计算随机积分的例子着手，引进并证明伊藤微分公式。

在前面三节中，我们从定义阶梯函数的随机积分开始，讨论了有界循序可测函数的随机积分以及局部有界循序可测过程的随机积分。

事实上，$\mathscr{P}_T$ 类中的函数

$$\mathscr{P}_T = \left\{H_t(\omega): P\left[\int_0^t H_s^2(\omega)ds < \infty, \forall t \geq 0\right] = 1\right\}$$

也可以定义随机积分。然而推导较前三节复杂，有兴趣的读者可参见文献[7]。

然而，光有随机积分的定义并没有解决随机积分如何计算的问题。正如在数学分析中，如果把 Riemann 积分作为积分和的极限来求，那么其复杂程度是可以想见的。只有建立了导数与原函数的一系列计算规则和牛顿公式之后，定积分计算才变得容易实施。

通过下面的例子，我们可以看到伊藤公式对于随机积分的计算相当于微积分中的牛顿公式。

在 3.5 节例 1 中，我们计算了 $\int_a^b W_t dW_t$。由该例可得

$$\int_0^t W_s dW_s = \frac{1}{2}W_t^2 - \frac{t}{2} \tag{3.6.1}$$

式(3.6.1)也可写成

$$\frac{1}{2}W_t^2 = \frac{t}{2} + \int_0^t W_s dW_s$$

或

$$\frac{1}{2}W_t^2 = \int_0^t \frac{1}{2}ds + \int_0^t W_s dW_s \tag{3.6.2}$$

从前几节的讨论中，我们已经知道：有界循序可测过程的积分(关于维纳过程的随机积分)是连续鞅，局部有界循序可测过程的积分是局部连续鞅。总而言之，随机积分的结果是随机过程。

如果把随机积分作为映照来看，那么 $W_t = \int_0^t dW_s$，即 W_t 的映照为 $\int_0^t dW_s$，而 $\frac{1}{2}W_t^2$ 的映照就不只是

$$\int_0^t W_s(\omega)d_sW_s$$

一项，而是式(3.6.2)，即它关于 dW_s 与 ds 的两个积分之和.

定义 3.6.1 设 W_t 是 $(\Omega, \mathscr{F}, P)$ 上的一维标准布朗运动，X_t 是 $(\Omega, \mathscr{F}, P)$ 上的随机过程，且

$$X_t = X_0 + \int_0^t u(s,\omega)ds + \int_0^t v(s,\omega)dW_s \qquad (3.6.3)$$

那么 X_t 也称为随机积分（或定义 3.6.1 形式的随机积分）.

这里随机积分 X_t 可以写成微分形式：

$$dX_t = udt + vdW_t \qquad (3.6.4)$$

因此式(3.6.2)可写成

$$d\left(\frac{1}{2}W_t^2\right) = \frac{1}{2}dt + W_tdW_t \qquad (3.6.5)$$

现在，我们叙述本节的主要结果：

定理 3.6.1 设 X_t 是式(3.6.4)形式的随机积分：

$$dX_t = udt + vdW_t$$

$g(t, x)$ 是 $[0, \infty) \times R \to R$ 的二阶连续可微函数，那么

$$Y_t = g(t, X_t) \qquad (3.6.6)$$

也是定义 3.6.1 形式的随机积分，且

$$dY_t = \frac{\partial}{\partial t}g(t, X_t)dt + \frac{\partial}{\partial x}g(t, X_t)dX_t + \frac{1}{2}\frac{\partial^2}{\partial x^2}g(t, X_t)(dX_t)^2 \qquad (3.6.7)$$

其中

$$dt \cdot dt = dt \cdot dW_t = dW_t \cdot dt = 0$$
$$dW_t \cdot dW_t = dt \qquad (3.6.8)$$

在证明定理 3.6.1（伊藤公式）之前，我们先看几个例子.

例 1 重新考虑 3.5 节例 1 中随机积分的计算:

$$I = \int_0^t W_t dW_t$$

设 $X_t = W_t$, $g(t, x) = \frac{1}{2}x^2$

那么

$$Y_t = g(t, W_t) = \frac{1}{2}W_t^2$$

利用伊藤公式

$$dY_t = \frac{\partial g}{\partial t}dt + \frac{\partial g}{\partial x}dW_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(dW_t)^2$$

$$= 0 + W_t dW_t + \frac{1}{2}(dW_t)^2$$

$$= W_t dW_t + \frac{1}{2}dt$$

因此

$$d\left(\frac{1}{2}W_t^2\right) = W_t dW_t + \frac{1}{2}dt$$

换言之,即

$$\frac{1}{2}W_t^2 = \int_0^t W_t dW_t + \frac{1}{2}t$$ ∎

例 2 计算随机积分

$$\int_0^t t dW_t$$

解 从经典的微积分知识可以得到启发, tW_t 这一项应该出现,故我们取

$$g(t, x) = t \cdot x$$

即

$$Y_t = g(t, W_t) = tW_t$$

根据伊藤公式

$$dY_t = W_t dt + t dW_t + 0$$

也就是

$$d(tW_t) = W_t dt + t dW_t$$

或

$$tW_t = \int_0^t W_s ds + \int_0^t s dW_s$$

即

$$\int_0^t s dW_s = tW_t - \int_0^t W_s ds \qquad (3.6.9)$$

式(3.6.9)与微积分中的分部积分公式一致。

伊藤公式的直观性证明：

把式(3.6.4)，(3.6.6)和(3.6.8)代入式(3.6.7)，得到

$$d(g(t, X_t)) = \frac{\partial g}{\partial t}(t, X_t)dt + \frac{\partial g}{\partial x}(t, X_t)[udt + vdW_t]$$

$$+ \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t, X_t)[udt + vdW_t]^2$$

$$= \left(\frac{\partial g}{\partial t} + u\frac{\partial g}{\partial x} + \frac{1}{2}v^2\frac{\partial^2 g}{\partial x^2}\right)dt$$

$$+ v\frac{\partial g}{\partial x}dW_t$$

即

$$g(t, X_t) = g(0, X_0) + \int_0^t \left(\frac{\partial g}{\partial s} + u\frac{\partial g}{\partial x}\right.$$

$$\left. + \frac{1}{2}v^2\frac{\partial^2 g}{\partial x^2}\right)ds + \int_0^t v\frac{\partial g}{\partial x}dW_s \qquad (3.6.10)$$

注意：式(3.6.10)是定义3.6.1意义下的随机积分。

我们看到为了证明式(3.6.10)，只要 u 和 v 是有界阶梯函

数就行。我们还可以假定 g、$\frac{\partial g}{\partial t}$、$\frac{\partial g}{\partial x}$ 和 $\frac{\partial^2 g}{\partial x^2}$ 都是有界函数。如果在这一情形下证明了公式(3.6.10)，那么我们可以在$[0,\infty)\times R$ 的一个紧子集上找到 $g_n\in C^2$ (二阶连续可微类)，使得 g_n、$\frac{\partial g_n}{\partial t}$、$\frac{\partial g_n}{\partial x}$ 和 $\frac{\partial^2 g_n}{\partial x^2}$ 对每一个 n 有界且分别一致收敛于 g、$\frac{\partial g}{\partial t}$、$\frac{\partial g}{\partial x}$ 和 $\frac{\partial^2 g}{\partial x^2}$。

利用泰勒展开，我们得

$$
\begin{aligned}
g(t, X_t) &= g(0, X_0) + \sum_j \Delta g(t_j, X_j) \\
&= g(0, X_0) + \sum_j \frac{\partial g}{\partial t}\Delta t_j + \sum_j \frac{\partial g}{\partial x}\Delta X_j \\
&\quad + \frac{1}{2}\sum_j \frac{\partial^2 g}{\partial t^2}\cdot(\Delta t_j)^2 \\
&\quad + \sum_j \frac{\partial^2 g}{\partial t\partial x}(\Delta t_j)(\Delta X_j) \\
&\quad + \frac{1}{2}\sum_j \frac{\partial^2 g}{\partial x^2}(\Delta X_j)^2 + \sum_j R_j
\end{aligned}
$$

其中

$$
\Delta t_j = t_{j+1} - t_j,\quad \Delta X_j = X_{t_{j+1}} - X_{t_j}
$$

$$
\Delta g(t_j, X_j) = g(t_{j+1}, X_{t_{j+1}}) - g(t_j, X_j)
$$

且

$$
R_j = 0(|\Delta t_j|^2 + |\Delta X_j|^2),\ \forall j
$$

如果 $\max_j \Delta t_j \to 0$，那么

$$
\sum_j \frac{\partial g}{\partial t}\Delta t_j = \sum_j \frac{\partial}{\partial t} g(t_j, X_j)\Delta t_j \xrightarrow{L_2} \int_0^t \frac{\partial g}{\partial t}(s, X_s)ds \tag{3.6.11}
$$

$$\sum_i \frac{\partial g}{\partial t} \Delta X_i = \sum_i \frac{\partial g}{\partial x}(t_i, X_i)\Delta X_i \xrightarrow{L_2}$$

$$\int_0^t \frac{\partial g}{\partial x}(s,\ X_s)dX_s \qquad (3.6.12)$$

更进一步

$$\sum_i \frac{\partial^2 g}{\partial x^2}(\Delta X_i)^2 = \sum_i \frac{\partial^2 g}{\partial x^2} u_i^2(\Delta t_i)^2$$

$$+ 2\sum_i \frac{\partial^2 g}{\partial x^2} u_i v_i (\Delta t_i)(\Delta W_i)$$

$$+ \sum_i \frac{\partial^2 g}{\partial x^2} v_i^2 (\Delta W_i)^2 \qquad (3.6.13)$$

当 $\max\limits_i \Delta t_i \to 0$ 时，式（3.6.13）右边的第一、二项趋于零.

例如

$$E\left[\left(\sum_i \frac{\partial^2 g}{\partial x^2} u_i v_i (\Delta t_i)(\Delta W_i)\right)^2\right]$$

$$= \sum_i E\left[\left(\frac{\partial^2 g}{\partial x^2} u_i v_i\right)^2\right](\Delta t_i)^3 \to 0, \text{ 当 } \Delta t_i \to 0$$

下面我们证明式（3.6.13）右边最末一项的极限：

$$\sum_i \frac{\partial^2 g}{\partial x^2} v_i^2 (\Delta W_i)^2 \xrightarrow[\Delta t_i \to 0]{L_2} \int_0^t \frac{\partial^2 g}{\partial x^2} v^2 ds \qquad (3.6.14)$$

为了证明式（3.6.14），令

$$a(t) = \frac{\partial^2 g}{\partial x^2} g(t,\ X_t) v^2(t,\ \omega)$$

$$a_i = a(t_i)$$

并考虑

$$E\left[\left(\sum_i a_i (\Delta W_i)^2 - \sum_i a_i \Delta t_i\right)^2\right]$$

$$= \sum_{i,j} E[a_i a_j((\Delta W_i)^2 - \Delta t_i)((\Delta W_j)^2 - \Delta t_j)]$$

$$= \left(\sum_{i<j} + \sum_{i=j} + \sum_{j<i}\right) E[a_i a_j \times ((\Delta W_i)^2 - \Delta t_i)((\Delta W_j)^2 - \Delta t_j)]$$

其中

$$\sum_{i<j} E[a_i a_j((\Delta W_i)^2 - \Delta t_i)((\Delta W_j)^2 - \Delta t_j)]$$
$$= E\sum_{i<j} a_i a_j((\Delta W_i)^2 - \Delta t_i) E((\Delta W_j)^2 - \Delta t_j)$$
$$= 0$$

这是因为 $a_i a_j((\Delta W_i)^2 - \Delta t_i)$ 与 $(\Delta W_j)^2 - \Delta t_j$ 独立。同理可得 $\sum_{j<i} = 0$。

又因为

$$\sum_i E[a_j^2((\Delta W_j)^2 - \Delta t_j)^2]$$
$$= \sum_i E(a_j^2) \cdot E[(\Delta W_j)^4 - 2(\Delta W_j)^2 \Delta t_j + (\Delta t_j)^2]$$
$$= \sum_i E a_j^2 \cdot [3(\Delta t_j)^2 - 2(\Delta t_j)^2 + (\Delta t_j)^2]$$
$$= 2\sum_i E a_j^2 \cdot (\Delta t_j)^2 \xrightarrow[\Delta t_j \to 0]{} 0$$

即

$$\sum_i a_j(\Delta W_j)^2 \xrightarrow{L_2} \int_0^t a(s) ds$$

我们常用

$$(dW_t)^2 = dt$$

表示 dW_t 与 dt 的关系。

当然，当 $\max_j \Delta t_j \to 0$ 时 $\sum R_j \to 0$。于是我们得到

$$g(t, X_t) = g(0, X_0) + \int_0^t \left(\frac{\partial g}{\partial s} + \frac{1}{2} v^2 \frac{\partial^2 g}{\partial x^2}\right) ds$$

$$+\int_0^t \frac{\partial g}{\partial x}\, dX_t$$

把 $dX_t = u dt + v dW_t$ 代入上式即得式(3.6.10). ∎

伊藤公式的严格证明可参阅文献[5, 7]。

习　题　3.6

1. 设 W_t 是布朗运动,试写出 W_t^2 的伊藤公式。

2. 设 W_t 是 ν 维布朗运动, $X_t = W_t + x$, $x \in R^\nu$, $f \in C^2$, 那么多维伊藤公式为

$$f(X_t) - f(x) = \sum_{k=1}^{\nu} \int_0^t \frac{\partial}{\partial x^k} f(X_s) dW_s^{(k)}$$

$$+ \frac{1}{2} \sum_{k=1}^{\nu} \int_0^t \frac{\partial^2}{(\partial x^k)^2} f(X_s) ds$$

设 $(W^{(1)}, W^{(2)})$ 是二维标准布朗运动,试写出 $W_t^{(1)} W_t^{(2)}$ 的伊藤公式。

3. 计算随机积分

$$\int_0^t s^2 dW_s$$

3.7　随机微分方程的解法

提要　本节首先引入白噪声过程的概念，并阐述白噪声过程与维纳过程的关系。然后讨论两个简单例子，说明随机微分方程的解法。最后给出多维伊藤公式及线性随机微分方程的解。

用下列形式表示的方程:

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t \tag{3.7.1}$$

称为随机微分方程，其中 X_t 是解过程; W_t 是标准布朗运动; $b(t, X_t)$, $\sigma(t, X_t)$ 是有界循序可测函数。

式(3.7.1)可写成积分方程的形式:

$$X_t = X_0 + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s \quad (3.7.2)$$

对于式(3.7.1)和(3.7.2),有两个问题需要解决:

(1) 在什么条件下,式(3.7.1)或(3.7.2)的解存在且唯一? 其解的性质如何?

(2) 怎样解式(3.7.1)微分方程?

现在我们先引入白噪声过程的概念。

设 $\{\xi_t\}_{t\in R}$ 是均值为零的广义平稳过程,其协方差函数为

$$E(\xi_{t+\tau}\xi_t) = s_0\delta(\tau) \quad (3.7.3)$$

其中 s_0 是常数。$\delta(\tau)$ 称为 Dirac-δ 函数,它具有如下性质:

$$\int_{-\infty}^{\infty} \delta(t)f(t)dt = f(0) \quad (3.7.4)$$

$f(\cdot)$ 是在 0 点连续的函数。

$$S(\nu) = s_0\int_{-\infty}^{\infty} \delta(\tau)e^{-i2\pi\nu\tau}d\tau = s_0 \quad (3.7.5)$$

称为 ξ_t 的谱密度函数。

定义 3.7.1 弱(广义)平稳过程 ξ_t,均值为零,谱密度等于常数,则称 ξ_t 为白噪声过程。

白噪声过程的功率谱密度函数等于常数,是说白噪声在各种频率上其功率是均匀的。这是白光的功率谱特点,因此称为白噪声。

考虑标准布朗运动

$\{W_t\}_{t\in R}$, $E(W_t - W_s)^2 = |t-s|$, $EW_t^2 = |t|$, $EW_s^2 = |s|$

因为
$$W_sW_t = \frac{1}{2}(W_t^2 + W_s^2) - \frac{1}{2}(W_t - W_s)^2$$

故

$$EW_tW_s = \frac{1}{2}[|t| + |s| - |t - s|] \tag{3.7.6}$$

现在定义

$$X_t = \frac{W_{t+\delta} - W_t}{\delta} \tag{3.7.7}$$

那么

$$\begin{aligned} E(X_tX_s) &= \frac{1}{\delta^2}[E(W_{t+\delta}W_{s+\delta}) + E(W_tW_s) \\ &\quad - E(W_{t+\delta}W_s) - E(W_tW_{s+\delta})] \\ &= \frac{1}{2\delta^2}[|t - s + \delta| + |t - s - \delta| - 2|t - s|] \\ &= \frac{1}{\delta}\max\left[0,\ 1 - \frac{|t - s|}{\delta}\right] \end{aligned} \tag{3.7.8}$$

于是，$\{X_t, -\infty < t < \infty\}$ 是广义平稳过程，且具有下列的功率谱密度：

$$\begin{aligned} S(\nu) &= \frac{1}{\delta}\int_{-\delta}^{\delta}\left(1 - \frac{|\tau|}{\delta}\right)e^{-i2\pi\nu\tau}d\tau \\ &= \frac{2}{\delta}\int_0^{\delta}\left(1 - \frac{\tau}{\delta}\right)\cos 2\pi\nu\tau\, d\tau \\ &= \frac{2}{2\pi\nu\delta}\left(1 - \frac{\tau}{\delta}\right)\sin 2\pi\nu\tau\Big|_{\tau=0}^{\delta} + \frac{2}{\delta 2\pi\nu\delta}\int_0^{\delta}\sin 2\pi\nu\tau\, d\tau \\ &= \frac{2(1 - \cos 2\pi\nu\delta)}{(2\pi\nu\delta)^2} = \left|\frac{\sin \pi\nu\delta}{\pi\nu\delta}\right|^2 \end{aligned} \tag{3.7.9}$$

从式(3.7.9)可以看出：当 $\delta \to 0$ 时 $S(\nu) \to 1$。所以，对于充分小的 δ，X_t 很像白噪声过程。我们把 $\{X_t\}$ 当 $\delta \to 0$ 时的形式极限看作白噪声，即我们把白噪声看作下列的形式极限(其极限的意义，仅仅指当 δ 充分小时，X_t 很像白噪声)：

$$\lim_{\delta \to 0}\frac{W_{t+\delta} - W_t}{\delta}$$

严格地讲，白噪声的功率无限大，它根本不是二阶矩随机过程。

有了白噪声的概念，我们可以详细分析本章开头提出来的两个随机微分方程问题了。

例 1 我们重新考虑人口生长模型：

$$\frac{dN_t}{dt} = a_t \cdot N_t, \quad N_0 \text{ 给定}$$

其中 $a_t = r_t + \alpha \cdot \xi_t$，$\xi_t$ 是白噪声，α 为常数。

为了简单起见，假定 $r_t = r$（常数）

$$\frac{dN_t}{dt} = \left(r + \alpha \frac{dW_t}{dt}\right) N_t$$

即

$$dN_t = rN_t dt + \alpha N_t dW_t \tag{3.7.10}$$

因此

$$\int_0^t \frac{1}{N_s} dN_s = rt + \alpha W_t \ (W_0 = 0) \tag{3.7.11}$$

为了计算式 (3.7.11) 左边的积分，令

$$g(t, x) = \ln x, \ x > 0$$

于是得到

$$\begin{aligned} d(\ln N_t) &= \frac{1}{N_t} dN_t + \frac{1}{2}\left(-\frac{1}{N_t^2}\right)(dN_t)^2 \\ &= \frac{dN_t}{N_t} - \frac{1}{2N_t^2}\alpha^2 N_t^2 dt \\ &= \frac{dN_t}{N_t} - \frac{1}{2}\alpha^2 dt \end{aligned}$$

因此

$$\frac{dN_t}{N_t} = d(\ln N_t) + \frac{1}{2}\alpha^2 dt$$

由式 (3.7.11) 得

$$\ln \frac{N_t}{N_0} = -\frac{1}{2}\alpha^2 \int_0^t ds + rt + \alpha W_t$$

$$= \left(r - \frac{1}{2}\alpha^2\right) t + \alpha W_t$$

即

$$N_t = N_0 \cdot \exp\left(\left(r - \frac{1}{2}\alpha^2\right) t + \alpha W_t\right) \quad (3.7.12)$$

式 (3.7.12) 是方程 (3.7.10) 的解。

例 2 我们回到本章开头的第二个问题，即

$$L\ddot{Q}_t + R\dot{Q}_t + \frac{1}{C} Q_t = F_t = G_t + \alpha\xi_t \quad (3.7.13)$$

这里 ξ_t 是均值为零的白噪声。

引入向量

$$X = X(t, \omega) = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} Q_t \\ \dot{Q}_t \end{bmatrix}$$

式 (3.7.13) 可写成

$$\begin{cases} \dot{X}_1 = X_2 \\ \dot{X}_2 = -\dfrac{R}{L} X_2 - \dfrac{1}{LC} X_1 + \dfrac{G_t}{L} + \dfrac{\alpha\xi_t}{L} \end{cases} \quad (3.7.14)$$

或写成矩阵形式

$$dX = AXdt + H_t dt + KdW_t \quad (3.7.15)$$

其中

$$dX = \begin{bmatrix} dX_1 \\ dX_2 \end{bmatrix}, A = \begin{bmatrix} 0 & 1 \\ -\dfrac{1}{CL} & -\dfrac{R}{L} \end{bmatrix}, H_t = \begin{bmatrix} 0 \\ \dfrac{1}{L} G_t \end{bmatrix}$$

$$K = \left[0, \frac{\alpha}{L}\right]^T$$

W_t 是一维布朗运动。

这样，我们就导出了二维的随机微分方程(3.7.15)。

为了求解随机微分方程(3.7.15)，我们需要用二维的伊藤微分公式。为此，我们给出下列定理。

定理 3.7.1（n 维伊藤公式） 设

$$dX_i = u_i dt + v_i dW_t,\ 1 \leqslant i \leqslant n \qquad (3.7.16)$$

g 是 $[0,\infty) \times R^n \to R$ 的 C^2 类函数（二阶连续可微函数）

$$Y_t = g(t, X_1, \cdots, X_n)$$

那么

$$dY_t = \frac{\partial g}{\partial t} dt + \sum_i \frac{\partial g}{\partial x_i} dX_i + \frac{1}{2} \sum_{i,j} \frac{\partial^2 g}{\partial x_i \partial x_j} (dX_i)(dX_j) \qquad (3.7.17)$$

证明略。

定理 3.7.2 设 $W_1, \cdots, W_m$ 是 m 个相互独立的一维布朗运动，$(W_1, \cdots, W_m)$ 是 m 维布朗运动。设

$$\begin{cases} dX_1 = u_1 dt + v_{11} dW_1 + \cdots + v_{1m} dW_m \\ \vdots \\ dX_n = u_n dt + v_{n1} dW_1 + \cdots + v_{nm} dW_m \end{cases} \qquad (3.7.18)$$

式(3.7.18)可写成矩阵形式：

$$dX = U dt + V dW$$

其中

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}, U = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}, V = \begin{bmatrix} v_{11} \cdots v_{1m} \\ \vdots \\ v_{n1} \cdots v_{nm} \end{bmatrix}$$

和

$$W = [W_1, \cdots, W_m]^T$$

再设 g 是 $[0, \infty) \times R^n \to R^k$ 的二阶连续可微函数，

$Y=[Y_1,\cdots,Y_k]^T$

$$Y=g(t,\ X)$$

那么一般的伊藤微分公式为

$$dY_r=\frac{\partial g_r}{\partial t}(t,X)dt+\sum_i\frac{\partial g_r}{\partial x_i}(t,X)dX_i$$
$$+\frac{1}{2}\sum_{i,j}\frac{\partial^2 g_r}{\partial x_i\partial x_j}(dX_i)(dX_j),1\leqslant r\leqslant k \quad (3.7.19)$$

其中 $(dW_i)(dW_j)=\delta_{ij}dt$，$dW_idt=dtdW_i=0$。这里 δ_{ij} 是 Kronecker-δ 函数,即

$$\delta_{ij}=\begin{cases}0, & 当\ i\neq j\\1, & 当\ i=j\end{cases}$$

这一定理的证明与最简单的伊藤微分公式证明类似，这里不赘述。

把定理 3.7.2 用于二维的向量值函数 $(g_1,g_2),h:[0,\infty)\times R^2\to R$，这里取

$$h(t,\ x_1,\ x_2)=\exp(-At)\begin{pmatrix}x_1\\x_2\end{pmatrix}$$

我们得到

$$d(\exp(-At)X)=-\exp(-At)AXdt+\exp(-At)dX \quad (3.7.20)$$

由式 (3.7.15) 可得

$$\exp(-At)dX-\exp(-At)AXdt$$
$$=\exp(-At)[H_tdt+KdW_t] \quad (3.7.21)$$

由式 (3.7.20) 与 (3.7.21) 得

$$d(X\cdot\exp(-At))=\exp(-At)H_tdt$$
$$+\exp(-At)KdW_t$$

即

$$\exp(-At)X-X_0=\int_0^t\exp(-As)H_sd_s$$

$$+\int_0^t \exp(-As)KdW_s \tag{3.7.22}$$

再令

$$l = \exp(-At)Kx, \quad X_t = W_t$$

利用伊藤公式可得

$$\int_0^t \exp(-As)KdW_s = \exp(-At)KW_t$$

$$+\int_0^t \exp(-At)AKW_t dt$$

于是式(3.7.22)可写成

$$X = \exp(At)\{X_0 + \exp(-At)KW_t$$

$$+\int_0^t [\exp(-As)H_s + \exp(-As)AKW_s]ds\}$$

$$\tag{3.7.23}$$

式(3.7.23)是方程(3.7.15)的解。

随机微分方程的求解，主要利用伊藤公式解出微分方程中的 X_t，X_t 可以用 W_t 的函数及其轨线关于 ds 的积分表示出来。

习 题 3.7

1.设

$$dX_t = -aX_t dt + \sigma dW_t$$

$$X_0 = c \in R$$

求随机微分方程的解。(X_t 称为 Ornstein-Uhlenbeck 速度过程)

3.8 随机微分方程解的存在性与唯一性

提要 本节阐明随机微分方程满足 Lipschitz 条件则其解存在且唯一。

设 $\sigma=(\sigma_{ij})$ 为矩阵，记 $|\sigma|^2=\sum_{i,j}|\sigma_{ij}|^2$；$b(x,t)=(b_1(x,t),\cdots,b_n(x,t))^T$，$\sigma(x,t)=(\sigma_{ij}(x,t))$，并假设 $b_i(x,t)$，$\sigma_{ij}(x,t)$ 均关于 $(x,t)\in R^n\times[0,T]$ 连续 $(T>0)$，而且 $\{W_t\}$ 是 n 维标准布朗运动，$\{\mathscr{F}_t\}$ 是它的自然 σ 代数族。如果 X_t 平方可积，并且具有下述分解：

$$X_t^i=\int_0^t b_i(X_s,s)ds+\sum_j\int_0^t\sigma_{ij}(X_s,s)dW_s^j$$
$$+x^i,\ \forall t\leqslant T\tag{3.8.1}$$

则我们称 $\{X_t\}$ 是满足初始条件

$$X_0=x_0,\ x_0\in R^n\tag{3.8.2}$$

的随机微分方程（X_t 是 t 的连续函数）

$$dX_t=b(X_t,t)dt+\sigma(X_t,t)dW_t\tag{3.8.3}$$

的一个强解。

定理 3.8.1 假定 $b(x,t)$，$\sigma(x,t)$ 是关于 $(x,t)\in R^n\times[0,T]$ 的连续向量值函数和连续矩阵值函数，且满足下列 Lipschitz 条件：

$$\begin{cases}|b(x,t)-b(\bar{x},t)|\leqslant c_*|x-\bar{x}|\\|\sigma(x,t)-\sigma(\bar{x},t)|\leqslant c_*|x-\bar{x}|\\|b(x,t)|\leqslant c(1+|x|)\\|\sigma(x,t)|\leqslant c(1+|x|)\end{cases}\tag{3.8.4}$$

这里 c_* 及 c 均为常数，那么对任何 $x\in R^n$，在 $(\Omega,\mathscr{F},\mathscr{F}_t,P)$ 中存在式 (3.8.2)，(3.8.3) 的唯一强解。此处总假定 W_t 是概率空间 $(\Omega,\mathscr{F},P)$ 中的布朗运动，$\mathscr{F}_t$ 为其自然 σ 代数族。

为了证明此定理，我们需要做一些准备。

引理 3.8.1 设非负可积函数 $\varphi(t)$ 满足条件

$$\varphi(t)\leqslant c\int_0^t\varphi(s)ds,\ \forall t\leqslant T$$

那么 $\varphi(t)=0$，对几乎所有 $t\in[0,\ T]$ 成立。

证明 我们不妨设 $c>0$。根据假设有

$$\begin{aligned}
\int_0^{\frac{1}{c}}\varphi(t)dt &\leqslant c\int_0^{\frac{1}{c}}\int_0^t\varphi(s)dsdt,\ \forall t<\frac{1}{c}\\
&= c\int_0^{\frac{1}{c}}\int_0^{\frac{1}{c}}\chi_{[s<t]}(s)\varphi(s)dsdt\\
&= c\int_0^{\frac{1}{c}}\int_0^{\frac{1}{c}}\chi_{[s<t]}(s,t)\varphi(s)dsdt\\
&= c\int_0^{\frac{1}{c}}\left(\int_0^{\frac{1}{c}}\chi_{[s<t]}(s,\ t)dt\right)\varphi(s)ds\\
&= c\int_0^{\frac{1}{c}}\left(\frac{1}{c}-s\right)\varphi(s)ds
\end{aligned}$$

由于 $\frac{1}{c}-s<\frac{1}{c}\left(\forall s\in\left(0,\ \frac{1}{c}\right]\right)$. 因此当 $\{\varphi(s)>0\}$ 的测度大于零时，将得出下述结论：

$$\int_0^{\frac{1}{c}}\varphi(s)dt<\int_0^{\frac{1}{c}}\varphi(t)dt$$

上式显然是错误的，故 $\left[0,\ \frac{1}{c}\right]\cap\{\varphi(t)>0\}$ 是零测集。再作代换 $\tilde{\varphi}(t)=\varphi\left(t+\frac{1}{c}\right)$，则

$$\tilde{\varphi}(t)\leqslant c\int_0^t\tilde{\varphi}(s)ds,\ \forall t\leqslant T-\frac{1}{c}$$

重复前面的步骤，推得 $\left[0,\ \frac{2}{c}\right]\cap\{\varphi(t)>0\}$ 是零测集。重复上面步骤，定理得证。 ∎

引理 3.8.2 设 H_t 是连续适应过程，H_0 有界，$E\int_0^T H_s^2ds<\infty$，于是 $\int_0^t H_sdW_s$ 是平方可积的连续鞅 $(\forall t\leqslant T)$.

证明 记 $H_t^{(n)}=\chi_{[0,T_n]}H_t$，此处 $T_n=\inf\{t,\ H_t\geqslant n\}$ 于是 $\int_0^t H_t^{(n)}dW_t$ 有定义，并且

$$\int_0^t H_t^{(n)}dW_t \xrightarrow[n\to\infty]{} \int_0^t H_t dW_t,\ \text{a.s.}$$

另一方面，由于

$$E\left|\int_0^t H_t^{(n)}dW_t-\int_0^t H_t^{(m)}dW_t\right|^2 \xrightarrow[n,m\to\infty]{} 0$$

故 $\int_0^t H_t^{(n)}dW_t$ 在 $L_2(\Omega)$ 中均方收敛，其极限 $\int_0^t H_t dW_t$ 是平方可积的。

现在继续证明定理 3.8.1：唯一性：假定 X_t 和 $\bar{X}_t$ 是两个解，则

$$X_t-\bar{X}_t=\int_0^t[b(X_s,\ s)-b(\bar{X}_s,\ s)]ds$$
$$+\int_0^t[\sigma(X_s,\ s)-\sigma(\bar{X}_s,\ s)]dW_s$$

于是根据 Lipschitz 条件 (3.8.4)，对于 $t<T$，有

$$E|X_t-\bar{X}_t|^2\leqslant 2E\left|\int_0^t(b(X_s,\ s)-b(\bar{X}_s,\ s)]ds\right|^2$$
$$+2E\int_0^t|\sigma(X_s,\ s)-\sigma(\bar{X}_s,\ s)|^2ds$$
$$\leqslant 2TE\int_0^t|b(X_s,\ s)-b(\bar{X}_s,\ s)|^2ds$$
$$+2E\int_0^t|\sigma(X_s,\ s)-\sigma(\bar{X}_s,\ s)|^2ds$$
$$\leqslant 2(1+T)c_*^2\int_0^t E|X_s-\bar{X}_s|^2ds$$

记 $\varphi(t)=E|X_t-\bar{X}_t|^2$，则

$$\varphi(t)\leqslant 2(1+T)c_*^2\int_0^t\varphi(s)ds$$

由引理 3.8.1，$\varphi(t)=0$，对于几乎所有 $t\leqslant T$ 成立。于是

$X_t = \overline{X}_t$, a.s.

存在性:

定义 $X_t^{(0)} = x$, 并定义

$$X_t^{(m+1)} = x + \int_0^t b(X_s^{(m)}, s)ds + \int_0^t \sigma(X_s^{(m)}, s)dW_s \tag{3.8.5}$$

于是

$$|X_t^{(1)} - X_t^{(0)}|^2 \leqslant 2\left|\int_0^t b(X_s^{(0)}, s)ds\right|^2 + 2\left|\int_0^t \sigma(X_s^{(0)}, s)dW_s\right|^2$$

记

$$M = 2c^2(T+1)(1+|x|)^2 + 2c_*^2(T+1)$$

则

$$E|X_t^{(1)} - X_t^{(0)}|^2 \leqslant 2c^2t^2(1+|x|)^2 + 2c^2t(1+|x|)^2 \leqslant Mt, \quad \forall t \leqslant T$$

应用数学归纳法: 现在假设 $m \leqslant k-1$, 下面的不等式成立:

$$E|X_t^{(m)} - X_t^{(m+1)}|^2 \leqslant \frac{(Mt)^{m+1}}{(m+1)!} \tag{3.8.6}$$

现在,我们来证明式(3.8.6)对 $m = k$ 仍然成立。

$$\begin{aligned}|X_t^{(k)} - X_t^{(k+1)}|^2 &\leqslant 2\left|\int_0^t [b(X_s^{(k-1)}, s) - b(X_s^{(k)}, s)]ds\right|^2 \\ &\quad + 2\left|\int_0^t [\sigma(X_s^{(k-1)}, s) - \sigma(X_s^{(k)}, s)]dW_s\right|^2\end{aligned} \tag{3.8.7}$$

两边取数学期望,并利用 Lipschitz 条件,得

$$E|X_t^{(k)} - X_t^{(k+1)}|^2 \leqslant 2tE\int_0^t |b(X_s^{(k-1)}, s) - b(X_s^{(k)}, s)|^2 ds$$

$$+2E\int_0^t|\sigma(X_s^{(k-1)},s)-\sigma(X_s^{(k)},s)|^2ds$$

$$\leqslant 2c_*^2tE\int_0^t|X_s^{(k-1)}-X_s^{(k)}|^2ds$$

$$+2c_*^2E\int_0^t|X_s^{(k-1)}-X_s^{(k)}|^2ds$$

$$\leqslant M\int_0^t E|X_s^{(k-1)}-X_s^{(k)}|^2ds$$

将式 (3.8.6) 代人上式得

$$E|X_t^{(k)}-X_t^{(k+1)}|^2\leqslant M\int_0^t\frac{(Ms)^k}{k!}ds$$

$$=\frac{(Ms)^{k+1}}{(k+1)!}$$

这样就证明了式 (3.8.6) 对所有 $m=0,1,2,\cdots$ 成立。

由式 (3.8.7)，我们还有

$$\sup_{0\leqslant t\leqslant T}|X_t^{(k)}-X_t^{(k+1)}|^2$$

$$\leqslant 2Tc_*^2\int_0^T|X_s^{(k-1)}-X_s^{(k)}|^2ds$$

$$+2\sup_{0\leqslant t\leqslant T}\left|\int_0^t[\sigma(X_s^{(k-1)},s)-\sigma(X_s^{(k)},s)]dW_s\right|^2 \tag{3.8.8}$$

由引理 3.8.2 和 Doob 不等式，得

$$E\sup_{0\leqslant t\leqslant T}\left|\int_0^t[\sigma(X_s^{(k-1)},s)-\sigma(X_s^{(k)},s)]dW_s\right|^2$$

$$\leqslant 2\sup_{0\leqslant t\leqslant T}E\left|\int_0^t[\sigma(X_s^{(k-1)},s)-\sigma(X_s^{(k)},s)]dW_s\right|^2$$

$$=2\sup_{0\leqslant t\leqslant T}E\int_0^t|\sigma(X_s^{(k-1)},s)-\sigma(X_s^{(k)},s)|^2ds$$

$$\leqslant 2c_*^2\int_0^T E|X_s^{(k-1)}-X_s^{(k)}|^2ds$$

因此

$$E \sup_{0\le t\le T} |X_t^{(k)} - X_t^{(k+1)}|^2$$

$$\leqslant 2Tc_*^2 \int_0^T E|X_s^{(k-1)} - X_s^{(k)}|^2 ds$$

$$+ 4c_*^2 \int_0^T E|X_s^{(k-1)} - X_s^{(k)}|^2 ds \qquad (3.8.9)$$

这样，从式(3.8.9)和(3.8.6)得

$$E \sup_{0\le t\le T} |X_t^{(m+1)} - X_t^{(m)}|^2 \leqslant K \frac{(MT)^m}{m!}$$

此处K是不依赖于m的常数，因此

$$P\left\{\sup_{0\le t\le T} |X_t^{(m+1)} - X_t^{(m)}| > \frac{1}{2^m}\right\} \leqslant 4^m K \frac{(MT)^m}{m!}$$

因为 $\sum[4^m(MT)^m/m!] < \infty$，所以从 Borel-Cantelli 引理可得，对几乎所有 ω，存在正整数 $m_0 = m_0(\omega)$ 使

$$\sup_{0\le t\le T} |X_t^{(m+1)}(\omega) - X_t^{(m)}(\omega)| \leqslant \frac{1}{2^m}, \quad \forall m \geqslant m_0(\omega)$$

从而部分和

$$X_t^{(k)}(\omega) = X_t^{(0)} + \sum_{m=0}^{k-1} (X_t^{(m+1)} - X_t^{(m)})$$

关于 $t\in[0, T]$ 一致收敛，$X_t^{(k)}$ 是连续过程，因此其极限仍为连续过程，记作 $X_t(\omega)$，从式(3.8.6)还可推出 $E|X_t|^2$ 是连续函数。

下面验证 $X_t(\omega)$ 满足随机微分方程(3.8.2)，(3.8.3)：由 Lipschitz 条件得

$$E\int_0^T |b(X_t^{(m)}, t) - b(X_t, t)|^2 dt \xrightarrow[m\to\infty]{} 0$$

$$E\int_0^T |\sigma(X_t^{(m)}, t) - \sigma(X_t, t)|^2 dt \xrightarrow[m\to\infty]{} 0$$

在式(3.8.5)中令 $m\to\infty$，即得式(3.8.3)。 ▮

如果我们将式(3.8.2)换成

$$x_0 = \xi \tag{3.8.2'}$$

此处 ξ 是与 W_t 独立的任一平方可积随机变量，定理仍成立．对此，我们只要平行地进行定理证明的各步骤，即可得出结论。此时 $\mathscr{F}_t = \sigma(\xi, W_s, s \leqslant t)$。

习 题 3.8

1．验证随机微分方程（n 维）

$$\begin{cases} dX_t = -\dfrac{1}{2} X_t dt + dW_t, \\ X_0 = x, \quad x \in R^n \end{cases}$$

有唯一解。

3.9 解过程的马氏性

提要 本节证明随机微分方程(3.8.2)，(3.8.3)的解是 3.1 节中所述的马尔可夫过程。

先回忆一下单调类定理．如果 $A, B \in \mathscr{B}$，则 $A \cap B \in \mathscr{B}$，称 $\mathscr{B}$ 为 π 族。

引理 3.9.1 设 $\mathscr{H}$ 为 Ω 上的实值函数的一个线性空间，如果

(1) $1 \in \mathscr{H}$；

(2) 设 $f_n \in \mathscr{H}$，$0 \leqslant f_n \uparrow f$，$f$ 有界，则 $f \in \mathscr{H}$；

(3) $\forall A \in \mathscr{B}$，有 $\chi_A \in \mathscr{H}$，这里 χ_A 表示 A 的示性函数，那么 $\mathscr{H}$ 包含 Ω 上一切 $\sigma(\mathscr{B})$ 可测的有界函数．

用此引理，可以证得：

引理 3.9.2 设 $\{\mathscr{F}_t\}$ 为递增 σ 代数族，X_t 是一个 $\{\mathscr{F}_t\}$ 适应过程．如果 Y 是有界随机变量，$\sigma(X_t)$ 表示一个随机变

量生成的 σ 代数，$Y\in\sigma(\mathscr{G},\sigma(X_t))$，此处 $\mathscr{G}$ 是与 $\mathscr{F}_t$ 独立的 σ 代数，那么

$$E(Y|\mathscr{F}_t)=E(Y|X_t),\ \text{a.s.} \tag{3.9.1}$$

证明 取集 $A\in\sigma(X_t)$，$B\in\mathscr{G}$，则

$$\begin{aligned}E[\chi_{A\cap B}|\mathscr{F}_t]&=E[\chi_A\chi_B|\mathscr{F}_t]\\&=\chi_AE[\chi_B|\mathscr{F}_t]\\&=P(B)\chi_A=\chi_AE[\chi_B|X_t]\\&=E[\chi_A\chi_B|X_t]=E[\chi_{A\cap B}|X_t]\end{aligned}\tag{3.9.2}$$

显然，形如 $A\cap B(A\in\sigma(X_t))$，$B\in\mathscr{G}$ 的集全体为一个 π 族，记为 $\mathscr{B}$. 记 $\mathscr{H}$ 为满足式 (3.9.1) 的有界随机变量全体. 显然 $\mathscr{H}$ 满足引理 3.9.1 的三个条件. 所以所有有界并且 $\sigma(\mathscr{B})=\sigma(\mathscr{G},\sigma(X_t))$ 可测的随机变量都属于 $\mathscr{H}$. 设 $\bar{\sigma}(\mathscr{G},\sigma(X_t))$ 是 $\sigma(\mathscr{B})$ 的完备化，那么式 (3.9.1) 对 $Y\in\bar{\sigma}(\mathscr{G},\sigma(X_t))$ 成立。 ∎

定理 3.9.1 随机微分方程 [(3.8.2)，(3.8.3)] 的解是马尔可夫过程。

证明 我们首先证明

$$E[\chi_A|\mathscr{F}_t]=E[\chi_A|X_t],\ \forall A\in\sigma(X_{t+\delta}),\delta>0 \tag{3.9.3}$$

包含式 (3.1.3)，即 X_t 是马尔可夫过程。

事实上，当 $A\in X_{t_n}$ 时，由于 $\mathscr{F}_{t_n}$ 是递增 σ 代数族，故由式 (3.9.3) 可推出 $(t_1<t_2<\cdots<t_n)$：

$$E[\chi_A|\mathscr{F}_{t_{n-1}}]=E[\chi_A|X_{t_{n-1}}]$$

又

$$E\{E[\chi_A|\mathscr{F}_{t_{n-1}}]|X_{t_1}\cdots X_{t_{n-1}}]\}$$

$$=E[\chi_A|X_{t_{n-1}}]=E[\chi_A|X_{t_1}\cdots X_{t_{n-1}}]$$

往证式(3.9.3)：如果记 $\overline{X}_\delta = X_{t+\delta}$，则

$$\begin{cases} d\overline{X}_\delta = b(\overline{X}_\delta,\ t+\delta)d\delta + d\overline{W}_\delta \\ \overline{X}_0 = X_t \end{cases}$$

其中 $\overline{W}_s = W_{t+s} - W_t$。由上节末，我们得知微分方程有唯一解 $\overline{X}_\delta \in \sigma(\sigma(X_t), \sigma(\overline{W}_s,\ \forall \leqslant \delta))$，取 $\mathscr{G} = \sigma(\overline{W}_s,\ \forall s \leqslant \delta)$，由于 $\mathscr{G}$ 与 $\mathscr{F}_t$ 独立.因此由引理 3.9.2 可推出式(3.9.3). ∎

习　　题　　3.9

1.设$X_t = \int_0^t H_s ds + W_t$，此处 W_t 为标准布朗运动，设 $0 = t_0^m < t_1^m < \cdots < t_{N(m)}^m = t$ 是 $[0,\ t]$ 的一列步长趋于零的分划，证明

$$\sum_{k=0}^{N(m)} (X_{t_{k+1}^m} - X_{t_k^m})^2 \xrightarrow[m\to\infty]{P} t$$

第三章　测　验　题

1.如果 W_t 是布朗运动，A 为正常数，证明：对任意 $T>0$ 有

$$Ee^{A W_t} \leqslant c,\ \forall 0 \leqslant t \leqslant T$$

其中 c 是依赖于 A，T 的正常数。

2.试写出随机积分

$$I_t = \int_0^t H_s dW_s$$

其中

$$H_s = \begin{cases} 1, & \text{当 } 0 \leqslant s < 1 \\ 2, & \text{当 } 1 \leqslant s < 2 \\ 3, & \text{当 } 2 \leqslant s \leqslant t,\ t > 2 \end{cases}$$

$\{W_t\}$ 是标准布朗运动，并证明 I_t 是连续鞅。

3.设 $\{W_t\}$ 是标准维纳过程，证明下列随机微分方程：

$$dx_1 = x_1 dW_t - \frac{1}{2}x_1 dt,\quad x_1(0) = 0$$

$$dx_2 = -x_1 dW_t - \frac{1}{2} x_2 dt, \quad x_2(0) = 1$$

的解是

$$x_1(t) = \sin W_t$$

$$x_2(t) = \cos W_t$$

4.试证明随机微分方程

$$dX_t = \frac{1}{3} dt + X_t \cos t dW_t$$

存在唯一解。

5.设 $M_t =$ 常数，若 $\{M_t\}_{t\geq 0}$ 既是弱平稳过程又是鞅，证明

$$P\{M_t = \text{常数}\} = 1$$

第四章　鞅和随机分析在控制理论中的某些应用

鞅论(包括随机分析——指随机积分与随机微分方程)在系统辨识、自适应控制、随机控制与滤波理论中的应用，是现代控制理论深入发展的自然结果，并非数学家们的故弄玄虚，下面的事实说明了这一点。

事实上，随机控制旨在研究各种随机过程以及由随机微分方程(或随机差分方程)所描述的系统的各种控制问题。滤波理论则考虑如何用被噪声污染了的输出数据，来估计随机系统在任意时刻 t 的状态。

在系统辨识方面，1976 年 L. Ljung 利用鞅收敛定理证明了，在持续激励条件下离散随机线性系统最小二乘辨识算法是收敛的。我国学者陈翰馥、郭雷获得了许多更精细的结果。

在自适应控制方面，1980 年 Goodwin-Ramadge 和 Caines 在证明随机自适应控制系统的输入、输出有界稳定性时又使用了几乎上鞅的收敛定理(见定理 3.8.1 和推论 3.8.1)。

由此可见，控制理论的许多领域均使用了鞅论和随机分析的思想和结论。目前，鞅论(包括随机分析)在控制理论中的应用还在向其深度和广度方面发展[21]。

在这一本近代概率引论的书中，我们不可能全面系统地介绍鞅论在控制理论各个方面的应用，只能介绍这方面应用的概貌以及一些基本事实和例子。许多最新的应用成果，读者可以在有关的最新文献中找到。

从概貌上说，鞅论(包括随机分析)，目前已经成功地应用于下列控制问题：

(1) 随机递推算法的收敛性分析，包括随机逼近算法的收敛性问题。

(2) 随机系统尤其是随机非线性系统的稳定性理论；

(3) 随机过程和随机系统的最优控制问题；

(4) 线性和非线性系统的滤波和预测问题；

(5) 复杂随机过程的扩散逼近等.

在本书中我们介绍四个方面的应用：4.1 节介绍鞅论在系统辨识中的应用；4.2 节给出鞅论在自适应控制中的应用；4.3 节介绍伊藤微分公式在随机控制中的应用；4.4 节讨论随机微分方程在滤波问题中的应用。

虽然这些例子都有一定的特殊性，但是从这些例子仍然可以看到鞅论(包括随机分析)在控制论中是十分有用的。

4.1 鞅收敛定理在系统辨识中的应用

提要 本节介绍推论 2.7.7 和 Kronecker 引理在系统辨识收敛性分析中的应用。

本节内容基本参照文献 [6] 的 3.2 节，为了阅读方便我们加了一些注释性文字。

设 $(\Omega, \mathscr{F}, P)$ 是基本概率空间. y_n 是 m 维输出，u_n 是 l 维输入，ε_n 是 m 维动态噪声. 我们考察动态系统中已知阶数 p, q 的多输入多输出离散时间系统：

$$y_n = A_1 y_{n-1} + \cdots + A_p y_{n-p} + B_1 u_{n-1} + \cdots + B_q u_{n-q} + \varepsilon_n \tag{4.1.1}$$

设 w_i 是 m' 维随机向量，记

$$\mathscr{F}_n = \sigma\{w_i,\ 0 \leqslant i \leqslant n\} \tag{4.1.2}$$

即 $\mathscr{F}_n$ 是由随机向量 $w_0, w_1, \cdots, w_n$ 生成的 σ 代数，显然有

$$\mathscr{F}_0 \subset \mathscr{F}_1 \subset \cdots \subset \mathscr{F}_n \subset \cdots \subset \mathscr{F}$$

并且设

$$E(w_n w_n^\tau | \mathscr{F}_{n-1}) = I, \quad E(w_n | \mathscr{F}_{n-1}) = 0 \tag{4.1.3}$$

其中 I 表示单位阵。

又设 F_n 是 $m \times m'$ 矩阵，它对于 $\mathscr{F}_{n-1}$ 可测，动态噪声表达为

$$\varepsilon_n = F_n w_n \tag{4.1.4}$$

应该强调，我们不要求 $E\operatorname{tr} F_n F_n' < \infty$，甚至于不要求 ε_n 的数学期望有穷，而只要求 w_n 满足式(4.1.3)。

假定输入 u_n 对于 $\mathscr{F}_n$ 可测，这样可以把线性和非线性的反馈控制包括在我们考察范围之内。

记

$$\theta^T = [A_1 \cdots A_p B_1 \cdots B_q] \tag{4.1.5}$$

$$\varphi_n^T = [y_n^T\ y_{n-1}^T \cdots y_{n-p+1}^T\ u_n^T \cdots u_{n-q+1}^T] \tag{4.1.6}$$

φ_{-1} 任意给定

那么方程(4.1.1)可以改写成

$$y_n = \theta^T \varphi_{n-1} + \varepsilon_n \tag{4.1.7}$$

记

$$Y_{n+1} = \begin{bmatrix} y_1^T \\ \vdots \\ y_{n+1}^T \end{bmatrix}, \quad \Psi_n = \begin{bmatrix} \varphi_0^T \\ \vdots \\ \varphi_n^T \end{bmatrix}, \quad H_n = \begin{bmatrix} \varepsilon_1^T \\ \vdots \\ \varepsilon_{n+1}^T \end{bmatrix} \tag{4.1.8}$$

从式(4.1.7)知

$$Y_{n+1} = \Psi_n \theta + H_n \tag{4.1.9}$$

这里模型集 $M(\theta)$ 由式(4.1.1)给出，未知参数 θ 由式(4.1.5)定义。

所谓系统辨识，就是根据输入输出数据 Ψ_n 和 Y_{n+1}，在一定的目标函数(准则函数)下，选一个模型使得它最好地拟合输入输出数据。

最直观而常用的目标函数是

$$J(\theta)=(Y_{n+1}-\Psi_n\theta)^T(Y_{n+1}-\Psi_n\theta)$$

也就是要选 θ_{n+1}，作为 θ 的估计，从而使

$$(Y_{n+1}-\Psi_n\theta_{n+1})^T(Y_{n+1}-\Psi_n\theta_{n+1})$$
$$=\min_\theta(Y_{n+1}-\Psi_n\theta)^T(Y_{n+1}-\Psi_n\theta) \quad (4.1.10)$$

即选 θ_{n+1} 使得

$$J(\theta_{n+1})=\min_\theta J(\theta)$$

设

$$\Psi_n^T\Psi_n>0 \text{ 即 } \sum_{i=0}^n\varphi_i\varphi_i^T>0 \quad (4.1.11)$$

那么很容易验算

$$(Y_{n+1}-\Psi_n\theta)^T(Y_{n+1}-\Psi_n\theta)=[\theta-(\Psi_n^T\Psi_n)^{-1}\Psi_n^TY_{n+1}]^T$$
$$\times\Psi_n^T\Psi_n[\theta-(\Psi_n^T\Psi_n)^{-1}\Psi_n^TY_{n+1}]$$
$$+Y_{n+1}^T[I-\Psi_n(\Psi_n^T\Psi_n)^{-1}\Psi_n^T]Y_{n+1} \quad (4.1.12)$$

在式(4.1.12)中依赖于 θ 的只有第一项，第一项等于零，式(4.1.12)左边达到最小值.因此使(4.1.12)式达到最小的 θ_{n+1} 为

$$\theta_{n+1}=(\Psi_n^T\Psi_n)^{-1}\Psi_n^TY_{n+1} \quad (4.1.13)$$

把式(4.1.8)代人可得

$$\theta_{n+1}=\left(\sum_{i=0}^n\varphi_i\varphi_i^T\right)^{-1}\sum_{i=0}^n\varphi_iy_{i+1}^T \quad (4.1.14)$$

式(4.1.13)和(4.1.14)都是最小二乘辨识的表达式。

在数理统计中，对于线性模型(4.1.9)，Ψ_n 通常称为设计矩阵.但那里的 Ψ_n 是确定性的，而现在是随机的.所以数理统

计中关于最小二乘估计的已有结论不能直接用到这里。

为了便于计算机在线运算，式(4.1.14)还可以写成递推形式。记

$$P_n = \left(\sum_{i=0}^{n-1} \varphi_i \varphi_i^T\right)^{-1}, \quad a_n = (1 + \varphi_n^T P_n \varphi_n)^{-1} \tag{4.1.15}$$

引理 4.1.1 最小二乘辨识满足下面的递推公式：

$$\theta_{n+1} = \theta_n + a_n P_n \varphi_n (y_{n+1}^T - \varphi_n^T \theta_n) \tag{4.1.16}$$

$$P_{n+1} = P_n - a_n P_n \varphi_n \varphi_n^T P_n \tag{4.1.17}$$

证明 我们先证一个一般的矩阵恒等式：

$$C - B^T(A + BC^{-1}B^T)^{-1}B = [C^{-1} + C^{-1}B^T A^{-1} B C^{-1}]^{-1} \tag{4.1.18}$$

由于

$$\begin{aligned}
&-(A+BC^{-1}B^T)^{-1}+A^{-1}-A^{-1}BC^{-1}B^T(A+BC^{-1}B^T)^{-1} \\
&= -(A + BC^{-1}B^T)^{-1} + [A^{-1}(A + BC^{-1}B^T) \\
&\quad - A^{-1}BC^{-1}B^T](A + BC^{-1}B^T)^{-1} \\
&= -(A + BC^{-1}B^T)^{-1} + (A + BC^{-1}B^T)^{-1} = 0
\end{aligned}$$

所以

$$\begin{aligned}
I &= I - C^{-1}B^T(A + BC^{-1}B^T)^{-1}B + C^{-1}B^T A^{-1}B \\
&\quad - C^{-1}B^T A^{-1} B C^{-1} B^T (A + BC^{-1}B^T)^{-1}B \\
&= (I + C^{-1}B^T A^{-1}B) - (C^{-1} + C^{-1}B^T A^{-1} B C^{-1})B^T \\
&\quad \times (A + BC^{-1}B^T)^{-1}B \\
&= (C^{-1} + C^{-1}B^T A^{-1} B C^{-1})C \\
&\quad - (C^{-1}+C^{-1}B^T A^{-1} BC^{-1})B^T(A+ BC^{-1}B^T)^{-1}B \\
&= (C^{-1}+C^{-1}B^T A^{-1}BC^{-1})[C - B^T(A+BC^{-1}B^T)^{-1}B]
\end{aligned}$$

于是得到式(4.1.18)。

根据 P_n 的定义

$$P_{n+1} = \left[\sum_{i=0}^{n-1} \varphi_i \varphi_i^T + \varphi_n \varphi_n^T\right]^{-1} = [P_n^{-1} + \varphi_n \varphi_n^T]^{-1} \tag{4.1.19}$$

在式(4.1.18)中，取 $C=P_n$，$B=\varphi_n^TP_n$，$A=1$，那么

$$
\begin{aligned}
[C^{-1}+C^{-1}B^TA^{-1}BC^{-1}]^{-1}&=[P_n^{-1}+\varphi_n\varphi_n^T]^{-1}\\
&=P_n-P_n\varphi_n(1+\varphi_n^TP_nP_n^{-1}P_n\varphi_n)^{-1}\varphi_n^TP_n\\
&=P_n-P_n\varphi_n(1+\varphi_n^TP_n\varphi_n)^{-1}\varphi_n^TP_n\\
&=P_n-a_nP_n\varphi_n\varphi_n^TP_n
\end{aligned}
\tag{4.1.20}
$$

从式(4.1.19)，(4.1.20)便可得到式(4.1.17)。

从式(4.1.14)和(4.1.17)可知

$$
\begin{aligned}
\theta_{n+1}&=P_{n+1}\sum_{i=0}^{n}\varphi_iy_{i+1}^T\\
&=(P_n-a_nP_n\varphi_n\varphi_n^TP_n)\left(\sum_{i=0}^{n-1}\varphi_iy_{i+1}^T+\varphi_ny_{n+1}^T\right)\\
&=\theta_n-a_nP_n\varphi_n\varphi_n^T\theta_n+P_n\varphi_ny_{n+1}^T\\
&\quad-a_nP_n\varphi_n\varphi_n^TP_n\varphi_ny_{n+1}^T\\
&=\theta_n-a_nP_n\varphi_n\varphi_n^T\theta_n+P_n\varphi_n(1-a_n\varphi_n^TP_n\varphi_n)y_{n+1}^T\\
&=\theta_n-a_nP_n\varphi_n\varphi_n^T\theta_n\\
&\quad+P_n\varphi_n[1-a_n(1+\varphi_n^TP_n\varphi_n-1)]y_{n+1}^T\\
&=\theta_n-a_nP_n\varphi_n\varphi_n^T\theta_n+a_nP_n\varphi_ny_{n+1}^T\\
&=\theta_n+a_nP_n\varphi_n(y_{n+1}^T-\varphi_n^T\theta_n)
\end{aligned}
$$

这就是式(4.1.16)。 ∎

我们把满足式(4.1.16)，(4.1.17)的递推估计称为递推最小二乘辨识（英文缩写为 RLS），并对它们赋予初值：θ_0 为任意给定的确定性矩阵，而

$$P_0=(mp+lq)I \tag{4.1.21}$$

这样就不会发生由于(4.1.11)式得不到满足所引起的麻烦。和引理 4.1.1 完全类似，利用矩阵恒等式(4.1.18)，容易验证

$$P_n=\left(\sum_{i=0}^{n-1}\varphi_i\varphi_i^T+\frac{1}{mp+lq}I\right)^{-1} \tag{4.1.22}$$

$$\theta_n = P_n \sum_{i=0}^{n-1} \varphi_i y_{i+1}^{\tau} + P_n P_0^{-1} \theta_0 \tag{4.1.23}$$

满足式(4.1.16)，(4.1.17)，并满足初值。

用 $\lambda_{\max}^n$ 和 $\lambda_{\min}^n$ 分别表示 P_n 的最大和最小特征值，并记

$$r_n \triangleq \operatorname{tr} P_n^{-1} = 1 + \sum_{i=0}^{n-1} \|\varphi_i\|^2 \tag{4.1.24}$$

下面用 k_i 表示不依赖于 $\omega \in \Omega$ 的常数。

r_n 单调非降，如果当 $n \to \infty$，r_n 有有限极限，那么非降矩阵列 $\sum_{i=0}^{n-1} \varphi_i \varphi_i^T$ 也有有限极限。由式(4.1.22)看出，$n \to \infty$ 时，P_n 收敛到常数阵 P。由式(4.1.23)知，当 $n \to \infty$ 时，如果 θ_n 也有极限，那么这种极限值中有一项为 $PP_0^{-1}\theta_0$，它依赖于初值的选取。所以对任意选择的 θ_0，式(4.1.23)就不可能给出一致估计。为了使对任意的初值 θ_0，θ_n 都是一致估计，必须要求

$$r_n \xrightarrow[n \to \infty]{} \infty, \quad \text{a.s.} \tag{4.1.25}$$

而 r_n 的极限有限是一种不足道的情形。

下面用数学分析中的一个定理作为引理。

引理 4.1.2 设 $a_i \geqslant 0$，$b_n = \sum_{i=0}^{n} a_i$，$b_n \xrightarrow[n \to \infty]{} \infty$，那么

$$\sum_{i=1}^{\infty} \frac{a_i}{b_i} = \infty, \quad \sum_{i=1}^{\infty} \frac{a_i}{b_i^{1+\delta}} < \infty, \quad \delta > 0 \tag{4.1.26}$$

设 A_i 是矩阵，若 $\sum_{i=1}^{\infty} A_i / b_i$ 收敛，则

$$\frac{1}{b_n} \sum_{i=1}^{n} A_i \xrightarrow[n \to \infty]{} 0 \tag{4.1.27}$$

证明 显然，对于固定的 n 有

$$\frac{a_{n+1}}{b_{n+1}}+\frac{a_{n+2}}{b_{n+2}}+\cdots+\frac{a_{n+r}}{b_{n+r}}$$

$$\geqslant\frac{b_{n+r}-b_n}{b_{n+r}}=1-\frac{b_n}{b_{n+r}}\xrightarrow[r\to\infty]{}1 \qquad (4.1.28)$$

所以对每一个 n，可取适当大的 r，使上述和大于 $\frac{1}{2}$，因此

$$\sum_{i=1}^{\infty}\frac{a_i}{b_i}=\infty$$

若 $x\geqslant 0$，$y\geqslant 0$，$0<\delta\leqslant 1$，则下面的不等式成立：

$$\delta x^{\delta-1}(x-y)\leqslant x^{\delta}-y^{\delta} \qquad (4.1.29)$$

因此对 $0<\delta\leqslant 1$（设 $b_0=b_1$），有

$$\sum_{i=1}^{\infty}\frac{a_i}{b_i^{1+\delta}}\leqslant\sum_{i=1}^{\infty}\frac{a_i}{b_ib_{i-1}^{\delta}}$$

$$=\sum_{i=1}^{\infty}\frac{\delta(b_i-b_{i-1})b_i^{\delta-1}}{\delta b_{i-1}^{\delta}b_i^{\delta}}$$

$$\leqslant\frac{1}{\delta}\sum_{i=1}^{\infty}\frac{b_i^{\delta}-b_{i-1}^{\delta}}{b_{i-1}^{\delta}b_i^{\delta}}$$

$$=\frac{1}{\delta}\sum_{i=1}^{\infty}\left(\frac{1}{b_{i-1}^{\delta}}-\frac{1}{b_i^{\delta}}\right)$$

$$=\frac{1}{\delta}\cdot\frac{1}{b_1^{\delta}}<\infty$$

当然对 $\delta>1$，更有 $\sum\limits_{i=1}^{\infty}\frac{a_i}{b_i^{1+\delta}}<\infty$。

再由引理 2.8.2（Kronecker 引理）知

$$\frac{1}{b_n}\sum_{i=1}^{n}A_i\xrightarrow[n\to\infty]{}0$$

▌

定理 4.1.1 设式(4.1.25)成立，而且存在常数 $k_0>0$，$\varepsilon\in[0,1)$，$\alpha\in\left[0,\ \frac{1}{2}\right)$ 以及有限随机变量 γ，使得

$$\operatorname{tr} F_n F_n^T \leqslant k_0 r_n^{\varepsilon},\ \text{a.s.},\ \forall n \tag{4.1.30}$$

$$\lambda_{\max}^n \leqslant \gamma / r_n^{1-\alpha},\ \alpha+\frac{\varepsilon}{2}<\frac{1}{2} \tag{4.1.31}$$

那么对于任给的 θ_0 以及式(4.1.21)给出的 P_0，由式(4.1.16)，(4.1.17)定出的最小二乘辨识是强一致的，并且收敛速度为

$$\|\theta_n-\theta\| = O(r_n^{\delta-\frac{1}{2}}),\ \forall\delta\in\left(\alpha+\frac{\varepsilon}{2},\ \frac{1}{2}\right] \tag{4.1.32}$$

最小二乘辨识是强一致的，这是指由式(4.1.16)和(4.1.17)定出的 θ_n 有

$$\theta_n \xrightarrow[n\to\infty]{} \theta,\ \text{a.s.}$$

其中 θ 为真参数。

证明 把式(4.1.7)和(4.1.4)代入式(4.1.23)得

$$\begin{aligned}\theta_n &= P_n\sum_{i=0}^{n-1}\varphi_i\varphi_i^T\theta + P_n\sum_{i=0}^{n-1}\varphi_i w_{i+1}^T F_{i+1}^T + P_n P_0^{-1}\theta_0 \\ &= \theta - \frac{1}{mp+lq}P_n\theta + P_n\sum_{i=0}^{n-1}\varphi_i w_{i+1}^T F_{i+1}^T + P_n P_0^{-1}\theta_0\end{aligned} \tag{4.1.33}$$

所以

$$\begin{aligned} r_n^{\frac{1}{2}-\delta}\|\theta_n-\theta\| \leqslant{} & \frac{1}{mp+lq} r_n^{\frac{1}{2}-\delta}\|P_n\|\|\theta\| \\ & + r_n^{\frac{1}{2}-\delta}\left\|P_n\sum_{i=0}^{n-1}\varphi_i w_{i+1}^T F_{i+1}^T\right\| + \|r_n^{\frac{1}{2}-\delta}P_n P_0^{-1}\theta_0\| \end{aligned} \tag{4.1.34}$$

由式(4.1.30)和(4.1.31)知

$$
r_n^{\frac{1}{2}-\delta}\|P_n\| \leqslant r_n^{\frac{1}{2}-\delta}\lambda_{\max}^n \leqslant \gamma r_n^{\frac{1}{2}-\delta+\alpha-1} = \gamma r_n^{\alpha-\frac{1}{2}-\delta} \xrightarrow[n\to\infty]{} 0 \tag{4.1.35}
$$

所以当 $n\to\infty$ 时，式(4.1.34)的右端的第一、三两项趋于零 a.s.。

设 a, b 为两个向量。矩阵 ab^T 的范数是 $ab^T(ab^T)^T$ 的最大特征值的平方根。矩阵 aa^T 的唯一非负特征值是 a^Ta，所以

$$
\|aa^T\| = \sqrt{a^Ta}
$$

$$
\|ab^T\| = \sqrt{b^\tau b}\ \sqrt{a^Ta} = \|a\|\|b\| \tag{4.1.36}
$$

注意到 r_i 及 F_i 关于 $\mathscr{F}_{i-1}$ 可测，由式(4.1.8)，(4.1.36)，(4.1.29)可得

$$
\begin{aligned}
E\left\|\frac{\varphi_{i-1}w_i^TF_i^T}{r_i^{\frac{1}{2}+\delta-\alpha}}\right\|^2 &= E\,\frac{\|\varphi_{i-1}\|^2\|F_iw_i\|^2}{r_i^{1+2\delta-2\alpha}} \\
&= E\,\frac{\|\varphi_{i-1}\|^2\operatorname{tr}F_iw_iw_i^TF_i^T}{r_i^{1+2\delta-2\alpha}} \\
&= E\left[\frac{\|\varphi_{i-1}\|^2}{r_i^{1+2\delta-2\alpha}}\cdot E(\operatorname{tr}F_iw_iw_i^TF_i^T|\mathscr{F}_{i-1})\right] \\
&= E\,\frac{\|\varphi_{i-1}\|^2\operatorname{tr}F_iF_i^T}{r_i^{1+2\delta-2\alpha}}
\end{aligned}
$$

由式(4.1.31)有

$$
\begin{aligned}
E\left\|\frac{\varphi_{i-1}w_i^TF_i^T}{r_i^{\frac{1}{2}+\delta-\alpha}}\right\|^2 &\leqslant k_0E\,\frac{\|\varphi_{i-1}\|^2r_i^{\varepsilon}}{r_i^{1+2\delta-2\alpha}} \\
&\leqslant k_0E\,\frac{\|\varphi_{i-1}\|^2}{r_i} \leqslant k_0
\end{aligned} \tag{4.1.37}
$$

由式(4.1.2)知

$$
E\left(\sum_{i=0}^{n}\frac{\varphi_{i-1}w_i^TF_i^T}{r_i^{\frac{1}{2}+\delta-\alpha}}\middle|\mathscr{F}_{n-1}\right)
$$

$$
= \sum_{i=0}^{n-1} \frac{\varphi_{i-1} w_i^T F_i^T}{r_i^{\frac{1}{2}+\delta-\alpha}} + E\left(\frac{\varphi_{n-1} w_n^T F_n^T}{r_n^{\frac{1}{2}+\delta-\alpha}} \middle| \mathscr{F}_{n-1} \right)
$$

$$
= \sum_{i=0}^{n-1} \frac{\varphi_{i-1} w_i^T F_i^T}{r_i^{\frac{1}{2}+\delta-\alpha}} + \frac{\varphi_{n-1}}{r_n^{\frac{1}{2}+\delta-\alpha}} E(w_n^T | \mathscr{F}_{n-1}) F_n^T
$$

$$
= \sum_{i=0}^{n-1} \frac{\varphi_{i-1} w_i^T F_i^T}{r_i^{\frac{1}{2}+\delta-\alpha}}
$$

所以

$$
\left(\sum_{i=0}^{n-1} \frac{\varphi_{i-1} w_i^T F_i^T}{r_i^{\frac{1}{2}+\delta-\alpha}}, \mathscr{F}_n \right) \tag{4.1.38}
$$

是二阶矩有限的鞅。由式(4.1.32)知

$$
\delta_0 \triangleq \delta - \alpha - \frac{\varepsilon}{2} > 0 \tag{4.1.39}
$$

和式(4.1.37)类似，可知

$$
\sum_{i=0}^{n} E\left[\left\| \frac{\varphi_{i-1} w_i^T F_i^T}{r_i^{\frac{1}{2}+\delta-\alpha}} \right\|^2 \middle| \mathscr{F}_{i-1} \right]
$$

$$
= \sum_{i=0}^{n} \frac{\|\varphi_{i-1}\|^2}{r_i^{1+2\delta-2\alpha}} \operatorname{tr} F_i F_i^T \leqslant k_0 \sum_{i=0}^{n} \frac{\|\varphi_{i-1}\|^2}{r_i^{1+2\delta_0}}
$$

由推论 2.7.7 知

$$
\sum_{i=1}^{\infty} \frac{\varphi_{i-1} w_i^T F_i^T}{r_i^{\frac{1}{2}+\delta-\alpha}} < \infty, \text{ a.s.}
$$

于是从式(4.1.27)便知

$$
\frac{1}{r_n^{\frac{1}{2}+\delta-\alpha}} \sum_{i=1}^{n} \varphi_{i-1} w_i^T F_i^T \xrightarrow[n\to\infty]{} 0, \text{ a.s.}
$$

注意到式(4.1.30)便知式(4.1.34)右端的第二项

$$
r_n^{\frac{1}{2}-\delta} \left\| P_n \sum_{i=0}^{n} \varphi_i w_{i+1}^T F_{i+1}^T \right\| \leqslant r_n^{\frac{1}{2}-\delta} \lambda_{\max}^n \left\| \sum_{i=0}^{n-1} \varphi_i w_{i+1}^T F_{i+1}^T \right\|
$$

$$
\leqslant \gamma r_n^{\frac{1}{2}-\delta} r_n^{\alpha-1} \left\| \sum_{i=0}^{n-1} \varphi_i w_{i+1}^T F_{i+1}^T \right\|
$$

$$
= \gamma r_n^{\alpha-\delta-\frac{1}{2}} \left\| \sum_{i=0}^{n-1} \varphi_i w_{i+1}^T F_{i+1}^T \right\| \xrightarrow[n \to \infty]{} 0, \quad \text{a.s.}
$$

因此从式(4.1.34),(4.1.35)可知式(4.1.32)成立，定理得证. ∎

从这个定理的证明可以看到：鞅收敛定理（这里用的是推论 2.7.7)在证明中起着重要的作用。由于随机系统中输出(序列)是相依的，这就决定了鞅论工具在系统辨识中重要地位。

4.2 几乎上鞅收敛定理在自适应控制中的应用

提要 本节介绍推论 2.8.1 在随机梯度自适应算法分析中的应用。

自适应控制是一个很大的领域，下面的例子仅仅是鞅论的一个应用，更多的应用可以在近代文献中找到[13]。

这里所讨论的离散随机系统可以用一个单输入单输出的 ARMAX 模型来描述，即

$$
\begin{aligned}
y_k + a_1 y_{k-1} + \cdots + a_n y_{k-n} = {} & b_0 u_{k-d} + b_1 u_{k-d-1} + \cdots \\
& + b_m u_{k-d-m} + w_k + c_1 w_{k-1} + \cdots + c_l w_{k-l}
\end{aligned} \tag{4.2.1}
$$

其中 $\{y_k\}$，$\{u_k\}$ 和 $\{w_k\}$ 分别表示输出、控制和噪声干扰。设 $\{w_k\}$ 是概率空间 $(\Omega, \mathscr{F}, P)$ 上的随机序列。设 x_0 为初始状态向量. $\mathscr{F}_0 = \sigma(x_0)$, $\mathscr{F}_t = \sigma(x_0, w_1, \cdots, w_t)$. 显然 $\forall\ t \geqslant 0, \mathscr{F}_0 \subset \mathscr{F}_t \subset \mathscr{F}, \mathscr{F}_t \uparrow$。假定 w_t 是鞅差序列，即

$$E\{w_t|\mathscr{F}_{t-1}\}=0,\ \text{a.s.},\ t\geqslant 1 \tag{4.2.2a}$$

$$E\{w_t w_t^T|\mathscr{F}_{t-1}\}=Q,\ \text{a.s.}$$
$$t\geqslant 1\ 且\ \mathrm{tr}Q<\infty \tag{4.2.2b}$$

$$\limsup_{N\to\infty}\frac{1}{N}\sum_{j=1}^{N}\|w_j\|^2<\infty,\ \text{a.s.} \tag{4.2.2c}$$

d 表示系统的时滞，本节讨论 $d=1$ 的情形。Goodwin-Ramadge-Caines（1981）的自适应算法为

$$\theta_k=\theta_{k-1}+\frac{\bar{a}}{r_{k-1}}\varphi_{k-1}[y_k-\varphi_{k-1}^T\theta_{k-1}],\ \bar{a}>0 \tag{4.2.3a}$$

$$r_{k-1}=1+\sum_{j=0}^{k-1}\varphi_j^T\varphi_j \tag{4.2.3b}$$

$$\varphi_k^T\theta_k=y_{k+1}^* \tag{4.2.3c}$$

其中

$$\varphi_{k-1}^T=(y_{k-1},\cdots,y_{k-n},u_{k-1},\cdots,u_{k-m-1},$$
$$-y_{k-1}^*,\cdots,-y_{k-l}^*) \tag{4.2.4}$$

从式（4.2.3c）中可以解自适应控制律。由于式（4.2.3c）中 θ_k 随外界环境和系统本身变化而变化，因此从式（4.2.3c）中解出的自适应控制律具有自适应性。

在自适应控制中要讨论的问题是：在什么条件下

$$(1)\ \lim_{N\to\infty}\frac{1}{N}\sum_{j=1}^{N}E\{(y_j-y_j^*)^2|\mathscr{F}_{j-1}\}=\gamma^2 \tag{4.2.5}$$

这里 γ^2 是因果反馈可能达到的最小均方误差。

$$(2)\ \limsup_{N\to\infty}\frac{1}{N}\sum_{j=1}^{N}y^2(j)<\infty \tag{4.2.6a}$$

$$\limsup_{N\to\infty}\frac{1}{N}\sum_{j=1}^{N}u^2(j)<\infty \tag{4.2.6b}$$

式（4.2.6a）和（4.2.6b）分别表示输入信号与输出信号的功

率均有限。

该问题的条件可以表示为下列假设：

假设 S：

(1) d 已知；

(2) m, n, l 的上界已知；

(3) $C(z^{-1})$ 和 $B(z^{-1})$ 所有零点在单位圆内。

为了获得式 (4.2.5) 和 (4.2.6) 的结果，我们先证明如下引理：

引理 4.2.1 如果假设 S 成立，且

$$C(z^{-1}) - \frac{\bar{a}}{2} \text{ 严格正实} \tag{4.2.7}$$

那么

$$\lim_{N\to\infty}\left(\frac{N}{r_N}\right)\frac{1}{N}\sum_{i=1}^{N}[\varepsilon_j - w_j]^2 = 0, \text{ a.s.} \tag{4.2.8}$$

其中

$$\varepsilon_j = y_j - \varphi_{j-1}^T\theta_{j-1}$$

证明 令

$$V_k = \tilde{\theta}_k^T\tilde{\theta}_k \tag{4.2.9}$$

其中 $\tilde{\theta}_k = \theta_k - \theta_0$，$\theta_0$ 表示系统的真参数。

由算法 (4.2.3a) 可得

$$\begin{aligned}
V_k &= \left(\tilde{\theta}_{k-1}^T + \frac{\bar{a}\varepsilon_k\varphi_{k-1}^T}{r_{k-1}}\right)\left(\tilde{\theta}_{k-1} + \frac{\bar{a}\varphi_{k-1}\varepsilon_k}{r_{k-1}}\right) \\
&= V_{k-1} + \frac{2\bar{a}\varepsilon_k\varphi_{k-1}^T}{r_{k-1}}\tilde{\theta}_{k-1} + \frac{\bar{a}^2\varepsilon_k^2\varphi_{k-1}^T\varphi_{k-1}}{r_{k-1}^2} \\
&= V_{k-1} + \frac{2\bar{a}}{r_{k-1}}\tilde{\theta}_{k-1}^T\varphi_{k-1}(\varepsilon_k - w_k) \\
&\quad + \frac{2\bar{a}}{r_{k-1}}\tilde{\theta}_{k-1}^T\varphi_{k-1}w_k + \frac{\bar{a}^2\varphi_{k-1}^T\varphi_{k-1}}{r_{k-1}^2}\varepsilon_k^2
\end{aligned}$$

因为

$$\varepsilon_k^2 = (\varepsilon_k - w_k)^2 + 2w_k(\varepsilon_k - w_k) + w_k^2 \quad (4.2.10)$$

令

$$b_{k-1} = -\tilde{\theta}_{k-1}^T \varphi_{k-1} \quad (4.2.11a)$$

$$z_{k-1} = \varepsilon_k - w_k \quad (4.2.11b)$$

于是

$$V_k = V_{k-1} - \frac{2\bar{a}}{r_{k-1}} b_{k-1} z_{k-1} - \frac{2\bar{a}}{r_{k-1}} b_{k-1} w_k + \frac{\bar{a}^2}{r_{k-1}^2} \varphi_{k-1}^T \varphi_{k-1} (z_{k-1}^2 + 2w_k z_{k-1} + w_k^2) \quad (4.2.12)$$

又因为

$$\begin{aligned} z_{k-1} &= y_k - \varphi_{k-1}^T \theta_{k-1} - w_k \\ &= -a_1 y_{k-1} - \cdots - a_n y_{k-n} + b_0 u_{k-1} + b_1 u_{k-2} + \cdots \\ &\quad + b_m u_{k-m-1} + c_1 w_{k-1} + \cdots \\ &\quad + c_l w_{k-l} - \varphi_{k-1}^T \theta_{k-1} \end{aligned} \quad (4.2.13)$$

因此 z_{k-1} 关于 $\mathscr{F}_{k-1}$ 可测。

由式 (4.2.12) 和 (4.2.13) 可得

$$\begin{aligned} &E\{V_k | \mathscr{F}_{k-1}\} \\ &= V_{k-1} - \frac{2\bar{a}}{r_{k-1}} b_{k-1} z_{k-1} + \frac{\bar{a}^2}{r_{k-1}^2} \varphi_{k-1}^T \varphi_{k-1} z_{k-1}^2 \\ &\quad + \frac{\bar{a}^2}{r_{k-1}^2} \varphi_{k-1}^T \varphi_{k-1} \cdot E\{w_k^2 | \mathscr{F}_{k-1}\} \\ &\leqslant V_{k-1} - \frac{2\bar{a}}{r_{k-1}} b_{k-1} z_{k-1} + \frac{\bar{a}^2}{r_{k-1}} z_{k-1}^2 \\ &\quad + \frac{\bar{a}^2}{r_{k-1}^2} \varphi_{k-1}^T \varphi_{k-1} \gamma^2 \end{aligned} \quad (4.2.14)$$

最后一个不等号成立，因为

(1) $\frac{\varphi_{k-1}^T\varphi_{k-1}}{r_{k-1}} \leqslant 1$;

(2) $E(w_k^2|\mathscr{F}_{k-1}) = \gamma^2$.

于是式(4.2.14)可以写成

$$E(V_k|\mathscr{F}_{k-1})$$

$$\leqslant V_{k-1} - \frac{2\bar{a}}{r_{k-1}}\left[b_{k-1} - \frac{\bar{a}+\rho}{2}z_{k-1}\right]z_{k-1}$$

$$-\rho\bar{a}\frac{z_{k-1}^2}{r_{k-1}} + \frac{\bar{a}^2}{r_{k-1}^2}\varphi_{k-1}^T\varphi_{k-1}\gamma^2 \qquad (4.2.15)$$

令

$$h_{k-1} = b_{k-1} - \frac{\bar{a}+\rho}{2}z_{k-1} \qquad (4.2.16)$$

这里 z_{k-1} 看作系统(4.2.16)的输入，h_{k-1} 看作该系统的输出。可以证明 $b_{k-1} = C(q^{-1})z_{k-1}$，其中

$$C(q^{-1}) = 1 + c_1q^{-1} + \cdots + c_lq^{-l}$$

q^{-1} 是后移算子，$q^{-1}z_k = z_{k-1}$。

$$h_{k-1} = C(q^{-1})z_{k-1} - \frac{\bar{a}+\rho}{2}z_{k-1}$$

$$= \left[C(q^{-1}) - \frac{\bar{a}+\rho}{2}\right]z_{k-1} \qquad (4.2.17)$$

由式(4.2.15)和(4.2.16)得

$$E(V_k|\mathscr{F}_{k-1}) \leqslant V_{k-1} - \frac{2\bar{a}}{r_{k-1}}h_{k-1}z_{k-1}$$

$$-\frac{\rho\bar{a}z_{k-1}^2}{r_{k-1}} + \frac{\bar{a}^2}{r_{k-1}^2}\varphi_{k-1}^T\varphi_{k-1}\gamma^2 \qquad (4.2.18)$$

因为 $C(z^{-1}) - \frac{\bar{a}}{2}$ 是严格正实的，所以存在一个小正数 $\rho>0$，使得 $C(z^{-1}) - \frac{\bar{a}+\rho}{2}$ 正实，再由式(4.2.17)得

$$2\bar{a}\sum_{j=1}^{k} h_{j-1}z_{j-1}+K>0$$

(关于这一点可以参见文献 [13] 定理 2.2.3)。

令

$$S_k=2\bar{a}\sum_{j=1}^{k} h_{j-1}z_{j-1}+K \tag{4.2.19}$$

定义一个非负随机变量序列

$$x_k=V_k+\frac{S_k}{r_{k-1}} \tag{4.2.20}$$

于是

$$\begin{aligned} E(x_k|\mathscr{F}_{k-1}) &= E(V_k|\mathscr{F}_{k-1})+\frac{S_k}{r_{k-1}} \\ &\leqslant V_{k-1}-\frac{2\bar{a}}{r_{k-1}}\ h_{k-1}z_{k-1}-\frac{\rho\bar{a}z_{k-1}^2}{r_{k-1}} \\ &\quad +\frac{\bar{a}^2}{r_{k-1}^2}\ \varphi_{k-1}^T\varphi_{k-1}\gamma^2+\frac{S_k}{r_{k-1}} \\ &\leqslant V_{k-1}+\frac{S_{k-1}}{r_{k-1}}-\frac{\rho\bar{a}z_{k-1}^2}{r_{k-1}} \\ &\quad +\frac{\bar{a}^2}{r_{k-1}^2}\ \varphi_{k-1}^T\varphi_{k-1}\gamma^2 \end{aligned} \tag{4.2.21}$$

式 (4.2.21) 可写成

$$E(x_k|\mathscr{F}_{k-1}) \leqslant x_{k-1}-\frac{\rho\bar{a}z_{k-1}^2}{r_{k-1}}+\frac{\bar{a}^2}{r_{k-1}^2}\ \varphi_{k-1}^T\varphi_{k-1}\gamma^2 \tag{4.2.22}$$

得到了式 (4.2.22) 后，我们就有可能套用几乎上鞅的收敛定理——推论 2.8.1，为此我们要检验以 (4.2.22) 不等式右端最后一项为通项的无穷级数是否收敛。

事实上

$$\sum_{k=1}^{N}\frac{\varphi_{k-1}^{T}\varphi_{k-1}}{r_{k-1}^{2}}\leqslant\sum_{k=1}^{N}\frac{\varphi_{k-1}^{T}\varphi_{k-1}}{r_{k-1}r_{k-2}}$$

$$=\sum_{k=1}^{N}\frac{r_{k-1}-r_{k-2}}{r_{k-1}r_{k-2}}$$

$$=\sum_{k=1}^{N}\left[\frac{1}{r_{k-2}}-\frac{1}{r_{k-1}}\right]\leqslant\frac{1}{r_{1}}<\infty$$

因此

$$\sum_{k=1}^{\infty}\frac{\varphi_{k-1}^{T}\varphi_{k-1}}{r_{k-1}^{2}}<+\infty \qquad (4.2.23)$$

由式(4.2.23)，根据推论 2.8.1，可以得到如下结论：

(1) $\lim\limits_{k\to\infty}x_k=x$, a.s. 且 $Ex<\infty$;

(2) $$\sum_{k=1}^{\infty}\frac{\rho\bar{a}z_{k-1}^{2}}{r_{k-1}}<+\infty,\ \text{a.s.} \qquad (4.2.24)$$

由式(4.2.24)再利用 Kronecker 引理，可得

$$\lim_{N\to\infty}\frac{1}{r_N}\sum_{k=1}^{N}z_k^2=0,\ \text{a.s.} \qquad (4.2.25)$$

欲获得所需要的结论，还要引用下列控制理论中的一个结果：

引理 4.2.2 如果系统(4.2.1)和(4.2.2)服从假设S的(3)，那么

$$\frac{1}{N}\sum_{k=1}^{N}u_k^2\leqslant\frac{K_1}{N}\sum_{k=0}^{N}y_k^2+K_2,\ \text{a.s.} \qquad (4.2.26)$$

其中 K_1, K_2 为常数。

定理 4.2.1 如果假设 S 中的(1)，(2)，(3)成立，且

$$C(z^{-1})-\frac{\bar{a}}{2},\ \text{严格正实}$$

对于系统 (4.2.1) 和 (4.2.2) 施行算法 (4.2.3)，那么

$$(1)\ \limsup_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N} y_k^2 < +\infty,\ \text{a.s.}; \qquad (4.2.27)$$

$$(2)\ \limsup_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N} u_k^2 < +\infty,\ \text{a.s.}; \qquad (4.2.28)$$

$$(3)\ \lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N} E[(y_k - y_k^*)^2 | \mathscr{F}_{k-1}] = \gamma^2. \qquad (4.2.29)$$

其中 γ^2 表示所有因果反馈所能产生的最小均方误差。

证明见文献 [13] 定理 6.5.1。值得注意的是：在这个定理证明中主要利用了引理 4.2.1 和引理 4.2.2。由此可见，几乎上鞅收敛定理——推论 2.8.1 在此证明中起着重要的作用。

4.3 随机最优控制问题

提要 本节考虑连续时间随机线性系统的最优控制问题。通过这一节，人们可以看到随机微分方程概念和伊藤微分公式在这个问题上的应用。

由于考虑非线性随机系统的最优控制是一件困难的事。为了说明随机微分方程和伊藤微分公式的应用，我们考虑用下列线性随机微分方程描述的系统：

$$dx(t) = Ax(t)dt + Bu(t)dt + dw(t) \qquad (4.3.1)$$

$$dy(t) = Cx(t)dt + de(t) \qquad (4.3.2)$$

其中 $x(t)$ 是 n 维状态向量，$u(t)$ 是 p 维控制向量，$y(t)$ 是 r 维输出向量。$\{w(t),\ t\in T\}$ 和 $\{e(t),\ t\in T\}$ 是独立的维纳过程，$Ew(t) = Ee(t) = 0$，且 $\mathrm{Cov}(dw(t)) = R_1dt$，$\mathrm{Cov}(de(t)) = R_2dt$。$A$，$B$，$C$，$D$，$R_1$ 和 R_2 是具有相应维数的矩阵，它们的元素可以是时间的按段连续的函数或常

数。

假定初始状态 $x(t_0)$ 是正态的，$Ex(t_0)=m$，$\mathrm{Cov}x(t)=R_0$。随机过程 $\{w(t),\ t\in T\}$ 和 $\{e(t),\ t\in T\}$ 独立于 $x(t_0)$。矩阵 R_0，R_1 是对称非负定矩阵，R_2 是对称正定矩阵。

控制变量的选择，应以一个准则函数为标准。这里准则函数选为期望损失

$$J=E\left\{x^T(t_1)Q_0x(t_1)+\int_{t_0}^{t}[x^T(t)Q_1x(t)+u^T(t)Q_2u(t)]\,dt\right\} \tag{4.3.3}$$

其中 Q_0 和 Q_1 是对称非负定矩阵，Q_2 是对称正定矩阵。

容许控制策略为控制信号在时刻 t 的值等于观测到 t 时刻为止的整个输出信号的函数。

我们考虑两种情形，即完全状态信息情形和不完全状态信息情形。完全状态信息意味着状态向量能够没有误差地测得。因为系统是由随机微分方程描述的，所以状态向量是马尔可夫过程，$x(t)$ 的未来状态在给定 $x(t)$ 与给定它过去值 $x(s)$，$s<t$ 的条件分布是相同的。因此在完全状态信息的情形下，容许控制策略 $u(t)$ 是 t 和 $x(t)$ 的函数，即

$$u(t)=\varphi(t,\ x(t))$$

在不完全状态信息的情形下，$u(t)$ 是 $\mathscr{Y}_t=\{y(s),\ t_0\leqslant s\leqslant t\}$ 的函数，随机控制问题可以完整地陈述如下：

考虑由随机微分方程 (4.3.1) 和 (4.3.2) 描述的系统。找出容许控制策略，使准则函数 (4.3.3) 极小。

为了解这个问题，我们将先证明几个有关的引理，然后导出有关期望损失的不等式。

引理 4.3.1 对于所有 x，y，函数 $l(x,\ y,\ u)$ 关于 u 有唯一的极小值，令 $u^0(x,\ y)$ 表示使 l 达到极小值的 u 值，

于是

$$\min_{u(x,y)} El(x, y, u) = El(x, y, u^0(x, y))$$

$$= E\min_u l(x, y, u) \qquad (4.3.4)$$

其中 x，y 为随机变量，u 为策略变量。

证明 对所有的容许策略，我们有

$$l(x, y, u) \geqslant l(x, y, u^0(x, y))$$

$$= \min_u l(x, y, u)$$

因此

$$El(x, y, u) \geqslant El(x, y, u^0(x, y))$$

$$= E \min_u l(x, y, u)$$

将左边关于所有允许策略取极小值得

$$\min_{u(x,y)} El(x, y, u) \geqslant El(x, y, u^0(x, y))$$

$$= E \min_u l(x, y, u) \qquad (4.3.5)$$

又因为 $u^0(x, y)$ 为一容许策略，我们有

$$El(x, y, u^0(x, y)) \geqslant \min_{u(x,y)} El(x, y, u) \qquad (4.3.6)$$

由式 (4.3.5) 和 (4.3.6) 可知，式 (4.3.4) 成立。

引理 4.3.2 假定 Riccati 方程

$$-\frac{dS}{dt} = A^TS + SA + Q_1 - SBQ_2^{-1}B^TS \qquad (4.3.7)$$

的初始条件为

$$S(t_1) = Q_0 \qquad (4.3.8)$$

式 (4.3.7) 在 $[t_0, t_1]$ 上有非负定解。设 x 是随机微分方程 (4.3.1) 的解，那么

$$x^T(t_1)Q_0x(t_1) + \int_{t_0}^{t_1}[x^T(t)Q_1x(t) + u^T(t)Q_2u(t)]dt$$

$$= x^T(t_0)S(t_0)x(t_0) + \int_{t_0}^{t_1}(u + Q_2^{-1}B^TSx)^TQ_2$$

$$\times (u + Q_2^{-1}B^TSx)dt + \int_{t_0}^{t_1} \mathrm{tr} R_1 S dt$$

$$+ \int_{t_0}^{t_1} dw^T(t)Sx + \int_{t_0}^{t_1} x^TSdw(t) \tag{4.3.9}$$

证明 由式(4.3.8)，我们有

$$x^T(t_1)Q_0x(t_1) = x^T(t_1)S(t_1)x(t_1)$$

$$= x^T(t_0)S(t_0)x(t_0) + \int_{t_0}^{t_1} d(x^TSx) \tag{4.3.10}$$

x 是随机微分方程(4.3.1)的解，因此，$d(x^TSx)$ 要利用伊藤微分公式来计算(参见 3.6 节)：

$$d(x^TSx) = (dx^T)Sx + x^TSdx + x^T \frac{dS}{dt} xdt$$

$$+ (\mathrm{tr}SR_1)dt \tag{4.3.11}$$

增量 dx 的协方差等于 R_1dt，由式(4.3.1)可得

$$x^TSdx = x^TS(Axdt + Budt + dw)$$

$$= [x^TSAx + x^TSBu]dt + x^TSdw$$

$$dx^TSx = [x^TA^TSx + u^TB^TSx]dt + dw^TSx$$

由式(4.3.7)得

$$x^T \frac{dS}{dt} xdt = [-x^TA^TSx - x^TSAx - x^TQ_1x$$

$$+ x^TSBQ_2^{-1}B^TSx]dt$$

于是式(4.3.11)可以写成

$$d(x^TSx) = [u^TB^TSx + x^TSBu - x^TQ_1x$$

$$+ x^TSBQ_2^{-1}B^TSx]dt + \mathrm{tr}(R_1S)dt$$

$$+ dw^TSx + x^TSdw$$

$$= [-u^TQ_2u - x^TQ_1x + (u+Q_2^{-1}B^TSx)^TQ_2(u$$

$$+ Q_2^{-1}B^TSx)]dt + \mathrm{tr}(R_1S)dt$$

$$+ dw^TSx + x^TSdw \tag{4.3.12}$$

上面最后一式是通过加一个减一个量 u^TQ_2u 得到的。重新

组织式 (4.3.12) 可以获得式 (4.3.9)。 ▮

利用引理 4.3.2，我们就能在不同情形下对系统 (4.3.1)，按照准则函数 (4.3.3) 解最优控制问题。

为了进行对比，我们首先考虑确定性的情形，此时 $w=0$。引理 4.3.2 给出了下列不等式：

$$
\begin{aligned}
x^T(t_1)Q_0x(t_1) &+ \int_{t_0}^{t_1}[x^T(t)Q_1x(t) + u^T(t)Q_2u(t)]dt \\
&= x^T(t_0)S(t_0)x(t_0) + \int_{t_0}^{t_1}[u + Q_2^{-1}B^TSx]^TQ_2[u \\
&\quad + Q_2^{-1}B^TSx]dt \geqslant x^T(t_0)S(t_0)x(t_0) \qquad (4.3.13)
\end{aligned}
$$

上式中的等号只有在控制策略为

$$
u = -Q_2^{-1}B^TSx = -Lx \qquad (4.3.14)
$$

时才得到。因为 Q_2 是正定的，故最优策略是唯一的。由引理 4.3.1，期望损失的极小值是

$$
\begin{aligned}
\min El &= Ex^T(t_0)S(t_0)x(t_0) \\
&= E(x(t_0) - m)^TS(t_0)(x(t_0) - m) \\
&\quad + Em^TS(t_0)x(t_0) \\
&\quad + Ex^T(t_0)S(t_0)m - Em^TS(t_0)m \\
&= E(x_0 - m)^TS(t_0)(x_0 - m) + m^TS(t_0)m \\
&= m^TS(t_0)m + \operatorname{tr}S(t_0)R_0 \qquad (4.3.15)
\end{aligned}
$$

其中 $x(t_0)$ 为正态随机变量；

$Ex(t_0) = m$；

$R_0 = E(x(t_0)-m)(x(t_0)-m)^T$。

对式 (4.3.9) 两边取数学期望，得

$$
\begin{aligned}
E\left\{x^T(t_1)Q_0x(t_1) + \int_{t_0}^{t_1}[x^T(t)Q_1x(t) + u^T(t)Q_2u(t)]dt\right\} \\
= E\left\{x^T(t_0)S(t_0)x(t_0) + \int_{t_0}^{t_1}[u + Q_2^{-1}B^TSx]^TQ_2[u\right.
\end{aligned}
$$

$$+ Q_2^{-1}B^TSx]dt + \int_{t_0}^{t_1}(\mathrm{tr}R_1S)dt\Big\}$$

$$\geqslant m^TS(t_0)m + \mathrm{tr}S(t_0)R_0 + \int_{t_0}^{t_1}(\mathrm{tr}R_1S)dt \quad (4.3.16)$$

式中等号在控制策略为

$$u = -Q_2^{-1}B^TSx = -Lx \quad (4.3.17)$$

时成立。

在完全状态信息的情形下，式 (4.3.17) 是允许控制策略. 因此，最优控制由式 (4.3.17) 给出，而损失函数的最小值是

$$\min El = m^TS(t_0)m + \mathrm{tr}S(t_0)R_0 + \int_{t_0}^{t_1}(\mathrm{tr}R_1S)dt \quad (4.3.18)$$

对于式 (4.3.9) 取数学期望并考察其极小值，有

$$\min E\left\{x^T(t_1)Q_0x(t_1) + \int_{t_0}^{t_1}[x^T(t)Q_1x(t) + u^T(t)Q_2u(t)]dt\right\}$$

$$= m^TS(t_0)m + \mathrm{tr}R_0S(t_0) + \int_{t_0}^{t_1}(\mathrm{tr}R_1S)dt$$

$$+ \min E\left\{\int_{t_0}^{t_1}(u + Lx)^TQ_2(u + Lx)dt\right\} \quad (4.3.19)$$

因为在给定 $\mathscr{F}_t = \sigma(y(S),\ t_0 \leqslant S \leqslant t)$的条件下 $x(t)$ 是正态分布的，其均值为 $\hat{x}$，协方差为 P，因此有

$$E\left[\int_{t_0}^{t_1}(u + Lx)^TQ_2(u + Lx)dt\right]$$

$$= E\int_{t_0}^{t_1}E[(u + Lx)^TQ_2(u + Lx)|\mathscr{F}_t]dt$$

$$= E\left\{\int_{t_0}^{t_1}(u + L\hat{x})^TQ_2(u + L\hat{x})dt\right.$$

$$\left. + \int_{t_0}^{t_1}(\mathrm{tr}L^TQ_2LP)dt\right\} \quad (4.3.20)$$

由于 P 不依赖于 u，有

$$E\left[x^T(t_1)Q_0x(t_1) + \int_{t_0}^{t_1}(x^TQ_1x + u^TQ_2u)dt\right]$$

$$\geqslant m^T S(t_0) m + \mathrm{tr} S(t_0) R_0 + \int_{t_0}^{t_1} (\mathrm{tr} R_1 S) dt$$

$$+ \int_{t_0}^{t_1} \mathrm{tr}(L^T Q_1 L P) dt \qquad (4.3.21)$$

上式中等号在控制策略为

$$u = -L\hat{x} = -LE[x(t) | \mathscr{F}_t] \qquad (4.3.22)$$

时取得。因而最优控制策略是 $x(t)$ 条件期望的线性函数。这里的 L 与式 (4.3.17) 中的 L 一致。

上面的结果可以归结为下面的定理：

定理 4.3.1 考虑由随机微分方程 (4.3.1) 和 (4.3.2) 描述的系统。假定 Riccati 方程 (4.3.7) 在 $t_0 \leqslant t \leqslant t_1$ 有解，那么控制律

$$u = -L\hat{x} = -LE[x(t) | \mathscr{F}_t]$$

使准则函数 (4.3.3) 取最小值，式中 L 由式 (4.3.14) 给定。期望损失的最小值为

$$\min El = m^T S(t_0) m + \mathrm{tr}(S(t_0) R_0) + \int_{t_0}^{t_1} (\mathrm{tr} S R_1) dt$$

$$+ \int_{t_0}^{t_1} (\mathrm{tr} S B Q_2^{-1} B^T S P) dt \qquad (4.3.23)$$

在引理 4.3.2 的证明中，我们使用了伊藤微分公式，而这里所考虑的系统是由随机微分方程描述的。

4.4 连续线性系统的最优估计——卡尔曼-布西滤波

提要 本节讨论连续时间系统状态 $x(t)$ 的最优估计问题，我们分三个子问题来叙述：(1) 问题的归结；(2) 证明使估计最优的充分必要条件是系统矩阵满足维纳-霍夫方程；(3) 最优滤波器设计。

在第三章的例3中，我们提到过滤波的问题，即设系统的方程为

$$L\ddot{Q}(t)+R\dot{Q}(t)+\frac{1}{C}Q(t)=F(t),\ Q(0)=Q_0$$

$$\dot{Q}(0)=I \tag{3.0.2}$$

$$Z(t)=Q(t)+e(t)$$

其中 $e(t)$ 是白噪声。

如何用量测到的 $z(s)$，$0\leqslant s\leqslant t$，估计 $Q(t)$ 和 $\dot{Q}(t)$，这就是滤波问题。

一般地，我们设系统的模型为

$$\dot{x}=F(t)x(t)+G(t)w(t) \tag{4.4.1}$$

$$z(t)=H(t)x+v(t) \tag{4.4.2}$$

其中 $t\geqslant t_0$，x 为 n 维向量，w 为 p 维向量，z 和 v 为 m 维向量，$F(t)$，$G(t)$，$H(t)$ 分别为连续的 $n\times n$，$n\times p$，$m\times n$ 矩阵(即矩阵的元素都是 t 的连续函数)，t 表示时间变量.

随机过程 $\{w(t),\ t\geqslant t_0\}$ 和 $\{v(t),\ t\geqslant t_0\}$ 是零均值高斯白噪声，其协方差矩阵分别为

$$E[w(t)w^T(s)]=Q(t)\delta(t-s)$$

和

$$E[v(t)v^T(s)]=R(t)\delta(t-s),\ (t,s\geqslant t_0)$$

其中 E 表示数学期望，δ 表示 Dirac-δ 函数。 $p\times p$ 矩阵 $Q(t)$ 对于任意 $t\geqslant t_0$ 为连续且非负定的，$m\times m$ 矩阵 $R(t)$ 对任意 $t\geqslant t_0$ 为连续且正定。

上述两个随机过程 $w(t)$ 和 $v(t)$ 假定彼此独立，于是对任意 $t,s\geqslant t_0$，有 $E[w(t)v^T(s)]=0$。其实也可以不受这一假设的限制，如果两个随机过程相关时，也可以推导出来．但有了这个假设可以简化许多运算。

初始状态 $x(t_0)$ 是 n 维零均值高斯随机向量，它与 $\{w(t),$

$t \geqslant t_0\}$ 和 $\{v(t),\ t \geqslant t_0\}$ 独立，且具有 $n \times n$ 协方差阵

$$E(x(t_0)x^T(t_0)) = P(t_0) \tag{4.4.3}$$

基于测量 $z(s)$，$(t_0 \leqslant s \leqslant t)$ 对状态 $x(t_1)$，$(t_1 \geqslant t_0)$ 所作的估计记作 $\hat{x}(t_1|t)$。

如果 $t_1 > t$，称 $\hat{x}(t_1|t)$ 为预测估计；如果 $t_1 = t$，称 $\hat{x}(t|t)$ 为滤波估计；如果 $t_1 < t$，则称 $\hat{x}(t_1|t)$ 为平滑估计.

$x(t_1)$ 的估计误差为

$$\tilde{x}(t_1|t) = x(t_1) - \hat{x}(t_1|t)$$

估计的性能指标是估计的均方误差，即

$$J[\tilde{x}(t_1|t)] = E[\tilde{x}^T(t_1|t)\tilde{x}(t_1|t)] \tag{4.4.4}$$

我们规定使 $J[\tilde{x}(t_1|t)]$ 最小的估计 $\hat{x}(t_1|t)$，取如下形式：

$$\hat{x}(t_1|t) = \int_{t_0}^{t} A(t_1, s)z(s)ds \tag{4.4.5}$$

其中 $A(t_1,\ s)$ 是关于 t_1，s 均连续可微的 $n \times m$ 矩阵.

式(4.4.5)的假定是受到离散时间最优估计的启发作出的(见文献[28]).

最优估计问题可叙述为给定量测 $\{z(s),\ t_0 \leqslant s \leqslant t\}$ 以后，求具有式(4.4.5)形式的状态 $x(t_1)$，$t_1 \geqslant t_0$ 的一个估计 $\hat{x}(t_1|t)$，使式(4.4.4)给出的均方误差最小.

式(4.4.5)中的 $A(t_1,\ s)$ 称为系统矩阵. 显然，这里的问题是求 $A(t_1,\ s)$，$t_0 \leqslant s \leqslant t$.

下面的定理以维纳-霍夫矩阵积分方程的形式给出了最优估计的必要与充分条件. 因此最优估计问题实际上是解特定的积分方程的问题.

定理 4.4.1 使 $\hat{x}(t_1|t)$ 为最优估计的充分必要条件是系统矩阵 $A(t_1,\ s)$ 应满足如下的维纳-霍夫方程，即对所有的 $t_0 \leqslant \sigma \leqslant t$，有

$$E[x(t_1)z^T(\sigma)] - \int_{t_0}^{t} A(t_1, s)E(z(s)z^T(\sigma))]ds = 0$$
(4.4.6)

证明 由估计误差定义及式(4.4.5),我们有

$$\tilde{x}(t_1|t) = x(t_1) - \hat{x}(t_1|t)$$
$$= x(t_1) - \int_{t_0}^{t} A^*(t_1,s)z(s)ds$$

其中 $A^*(t,s)$ 是任意系统矩阵,误差 $\tilde{x}(t_1|t)$ 的协方差矩阵为

$$E[\tilde{x}(t_1|t)\tilde{x}^T(t_1|t)]$$
$$= E\left\{\left[x(t_1) - \int_{t_0}^{t} A^*(t_1, s)z(s)ds\right]\left[x(t_1) - \int_{t_0}^{t} A^*(t_1, \sigma)z(\sigma)d\sigma\right]^T\right\}$$
$$= E[x(t_1)x^T(t_1)] - \int_{t_0}^{t} A^*(t_1,s)E[z(s)x^T(t_1)]ds$$
$$- \int_{t_0}^{t} E[x(t_1)z^T(\sigma)]A^{*T}(t_1, \sigma)d\sigma$$
$$+ \int_{t_0}^{t}\int_{t_0}^{t} A^*(t_1, s)E[z(s)z^T(\sigma)]A^{*T}(t_1,\sigma)dsd\sigma$$
(4.4.7)

而均方误差是协方差矩阵之迹.因为式(4.4.7)右边第三项是第二项的转置,所以两项和的迹就是第三项之两倍. 因此,可以把均方误差表示为

$$J[\tilde{x}(t_1|t)] = \mathrm{tr}\{E[x(t_1)x^T(t_1)]$$
$$- 2\int_{t_0}^{t} E[x(t_1)z^T(\sigma)]A^{*T}(t_1, \sigma)d\sigma$$
$$+ \int_{t_0}^{t} A^*(t_1, s)ds\int_{t_0}^{t} E[z(s)z^T(\sigma)]$$
$$\times A^*(t_1, \sigma)d\sigma \qquad (4.4.8)$$

方程(4.4.8)表明了均方误差是 $A^*(t, s)$ 的函数。

现在假设存在一个系统矩阵 $A(t_1, s)$，它使性能指标(4.4.8)最小，另外构造一个系统矩阵 $A^*(t,s)$

$$A^*(t_1, s) = A(t_1, s) + \varepsilon A_\varepsilon(t_1, s) \qquad (4.4.9)$$

其中 $A_\varepsilon(t_1, s)$ 是任意的系统矩阵，ε 为一个标量参数，$\varepsilon A_\varepsilon(t_1, s)$ 称为 $A(t_1, s)$ 的变分。

如果把式（4.4.9）代入式（4.4.8），则对于某固定的 $A_\varepsilon(t, s)$ 来说，$J[\tilde{x}(t_1|t)]$ 就变成了 ε 的函数。因此，如果 $A(t_1, s)$ 是使性能指标为最小的系统矩阵，那么只有当 $\varepsilon = 0$ 时，性能指标才能最小。也就是说，性能指标关于 ε 的偏导数，当 $\varepsilon = 0$ 时必然为零。这就得到了 $A(t_1, s)$ 的必要条件。推导出这个必要条件后，还可证明它是充分的。

把式(4.4.9)代入式(4.4.8)得

$$\begin{aligned}
J[\tilde{x}(t_1|t)] = \operatorname{tr}\{&E[x(t_1)x^T(t_1)] \\
&- 2\int_{t_0}^{t} E[x(t_1)z^T(\sigma)][A(t_1, \sigma) \\
&+ \varepsilon A_\varepsilon(t_1,\sigma)]^T d\sigma + \int_{t_0}^{t}[A(t_1, s) \\
&+ \varepsilon A_\varepsilon(t_1,s)]ds\int_{t_0}^{t} E[z(s)z^T(\sigma)] \\
&\times [A(t_1, \sigma) + \varepsilon A_\varepsilon(t_1, \sigma)]d\sigma\} \qquad (4.4.10)
\end{aligned}$$

于是

$$\begin{aligned}
\frac{\partial J}{\partial \varepsilon} = \operatorname{tr}\Big\{&-2\int_{t_0}^{t} E[x(t_1)z^T(\sigma)]A_\varepsilon(t_1, \sigma)^T d\sigma \\
&+ \int_{t_0}^{t}[A(t_1, s) + \varepsilon A_\varepsilon(t_1, s)]ds \\
&\times \int_{t_0}^{t} E[z(s)z^T(\sigma)]A_\varepsilon^T(t_1, \sigma)d\sigma \\
&+ \int_{t_0}^{t} A_\varepsilon(t_1,s)ds\int_{t_0}^{t} E(z(s)z^T(\sigma)][A(t_1,\sigma)
\end{aligned}$$

$$+ \varepsilon A_\varepsilon(t_1,\sigma)]^T d\sigma\Big\}$$

因为上式右边最后一项是第二项的转置，故这两项和的迹是最后一项迹的两倍，所以

$$\frac{\partial J}{\partial \varepsilon} = 2\mathrm{tr}\left\{\int_{t_0}^{t}[A(t_1,s) + \varepsilon A_\varepsilon(t_1,s)]ds \times \int_{t_0}^{t} E[z(s)z^T(\sigma)]A_\varepsilon^T(t_1,\sigma)d\sigma - \int_{t_0}^{t} E[x(t_1)z^T(\sigma)]A_\varepsilon^T(t_1,\sigma)d\sigma\right\}$$

令 $\dfrac{\partial J}{\partial \varepsilon} = 0$，可得

$$\mathrm{tr}\left\{\int_{t_0}^{t} A(t_1,s)ds\int_{t_0}^{t} E(z(s)z^T(\sigma))A_\varepsilon^T(t_1,\sigma)d\sigma - \int_{t_0}^{t} E[x(t_1)z^T(\sigma)]A_\varepsilon^T(t_1,\sigma)d\sigma\right\} = 0$$

即

$$\mathrm{tr}\left[\int_{t_0}^{t_1}\left\{E(x(t_1)z^T(\sigma)) - \int_{t_0}^{t_1} A(t_1,s)E[z(s)z^T(\sigma)]ds\right\} \times A_\varepsilon^T(t_1,\sigma)d\sigma\right] = 0 \qquad (4.4.11)$$

由于式(4.4.11)对任意系数矩阵 $A_\varepsilon(t_1, s)$ 都成立，因此 $A(t_1, s)$ 必须满足维纳-霍夫方程，即

$$E[x(t_1)z^T(\sigma)] - \int_{t_0}^{t} A(t_1,s)E[z(s)z^T(\sigma)]ds = 0 \qquad (4.4.6)$$
$$\forall t_0 \leqslant \sigma \leqslant t$$

往证。充分性：把式(4.4.10)展开，得

$$J[\tilde{x}(t_1|t)] = \mathrm{tr}\left\{E[x(t_1)x^T(t_1)] - z\int_{t_0}^{t} E[x(t_1)z^T(\sigma)]A^T(t_1,\sigma)d\sigma - z\varepsilon\int_{t_0}^{t} E[x(t_1)z^T(\sigma)]A_\varepsilon^T(t_1,\sigma)d\sigma\right.$$

$$
\begin{aligned}
&+\int_{t_0}^{t} A(t_1,s)ds\int_{t_0}^{t} E[z(s)z^T(\sigma)]A^T(t_1,\sigma)d\sigma \\
&+\varepsilon\int_{t_0}^{t} A_\varepsilon(t_1,s)ds\int_{t_0}^{t} E[z(s)z^T(\sigma)]A^T(t_1,\sigma)d\sigma \\
&+\varepsilon\int_{t_0}^{t} A(t_1,s)ds\int_{t_0}^{t} E[z(s)z^T(\sigma)]A_\varepsilon^T(t_1,\sigma)d\sigma \\
&+\varepsilon^2\int_{t_0}^{t} A_\varepsilon(t_1,s)ds\int_{t_0}^{t} E[z(s)z^T(\sigma)]A_\varepsilon^T(t_1,\sigma)d\sigma\Big\}
\end{aligned}
\tag{4.4.12}
$$

把上式右端的第一、二、四项与式(4.4.8)对照，发现：这三项等于系统矩阵为 $A(t_1,s)$ 时的均方误差，即最小性能指标，我们用 $J^0[\tilde{x}(t_1|t)]$ 表示这些项。

又因为式(4.4.12)第五项是第六项的转置，故两者和的迹等于第六项迹的两倍。这样式(4.4.12)可简化为

$$
\begin{aligned}
J[\tilde{x}(t_1|t)] = {} & J^0[\tilde{x}(t_1|t)] \\
& - 2\varepsilon\,\mathrm{tr}\Big\{\Big[\int_{t_0}^{t}\Big\{E[x(t_1)z^T(\sigma)] \\
& - \int_{t_0}^{t} A(t_1,s)E[z(s)z^T(\sigma)]ds\Big\}A_\varepsilon^T(t_1,\sigma)d\sigma\Big]\Big\} \\
& + \varepsilon^2\mathrm{tr}\Big\{\int_{t_0}^{t}\int_{t_0}^{t} A_\varepsilon(t_1,s)E[z(s)z^T(\sigma)]A_\varepsilon^T(t_1,\sigma)d\sigma ds\Big\}
\end{aligned}
\tag{4.4.13}
$$

由式(4.4.2)知

$$
\begin{aligned}
&E[z(s)z^T(\sigma)] \\
&\quad = H(s)E[x(s)x^T(\sigma)]H^T(\sigma) + H(s)E[x(s)v^T(\sigma)] \\
&\qquad + E[v(s)x^T(\sigma)]H^T(\sigma) + R(s)\delta(s-\sigma)
\end{aligned}
$$

上式右边头三项是非负定的，最后一项 $R(s)\delta(s-\sigma)$ 是正定的[因为 $R(s)$ 正定]。由此可见式(4.4.13)右边最后一项是正的。又因为 $A(t_1,s)$ 满足式(4.4.6)，式(4.4.13)右边第二项等于零，因此

$$J[\tilde{x}(t_1|t)] > J^0[\tilde{x}(t_1|t)]$$

这就是说，如果维纳-霍夫方程成立，则 $A(t_1, s)$ 使性能指标达到最小。 ∎

因为

$$\tilde{x}(t_1|t) = x(t_1) - \int_{t_0}^{t} A(t_1, s) z(s) ds$$

维纳-霍夫方程可以写成

$$E\left\{\left[x(t_1) - \int_{t_0}^{t} A(t_1, s) z(s) ds\right] z^T(\sigma)\right\} = 0$$

即

$$E[\tilde{x}(t_1|t) z^T(\sigma)] = 0 \quad (t_0 \leqslant \sigma \leqslant t) \tag{4.4.14}$$

式 (4.4.14) 是维纳-霍夫方程的另一种形式。

推论 4.4.1 $E[\tilde{x}(t_1|t)\hat{x}^T(t_1|t)] = 0$ (4.4.15)

证明

$$\begin{aligned} & E[\tilde{x}(t_1|t)\hat{x}^T(t_1|t)] \\ & = E\left\{\tilde{x}(t_1|t)\left[\int_{t_0}^{t} A(t_1, s) z(s) ds\right]^T\right\} \\ & = \int_{t_0}^{t} E[\tilde{x}(t_1|t) z^T(s)] A^T(t_1, s) ds \end{aligned}$$

由式 (4.4.14) 可得式 (4.4.15)。 ∎

引理 4.4.1 设 $\hat{x}(t_1|t)$ 和 $\hat{x}_*(t_1|t)$ 都是 $x(t_1)$ 的最优估计，那么

$$E\{[\hat{x}(t_1|t) - \hat{x}_*(t_1|t)][\hat{x}(t_1|t) - \hat{x}_*(t_1|t)]^T\} = 0 \tag{4.4.16}$$

证明 令

$$\tilde{x}(t_1|t) = x(t_1) - \hat{x}(t_1|t)$$
$$\tilde{x}_*(t_1|t) = x(t_1) - \hat{x}_*(t_1|t)$$

由推论 4.4.1 知

$$E[\tilde{x}(t_1|t)\hat{x}^T(t_1|t)] = 0 \tag{4.4.17}$$

$$E[\tilde{x}(t_1|t)\hat{x}_*^T(t_1|t)]=0 \tag{4.4.18}$$

从式(4.4.17)减去式(4.4.18)得

$$E\{\tilde{x}(t_1|t)[\hat{x}(t_1|t)-\hat{x}_*(t_1|t)]^T\}=0 \tag{4.4.19}$$

同理可得

$$E\{\tilde{x}_*(t_1|t)[\hat{x}(t_1|t)-\hat{x}_*(t_1|t)]^T\}=0 \tag{4.4.20}$$

由式(4.4.20)减去式(4.4.19)得

$$E\{[\tilde{x}_*(t_1|t)-\tilde{x}(t_1|t)][\hat{x}(t_1|t)-\hat{x}_*(t_1|t)]^T\}=0$$

因为

$$\tilde{x}_*(t_1|t)-\tilde{x}(t_1|t)=\hat{x}(t_1|t)-\hat{x}_*(t_1|t)$$

因此引理得证。 ▮

上面推导了维纳-霍夫方程和有关的命题，有了这些准备我们可以来推导最优滤波了。

这里假定 $\{w(t),\ t\geq t_0\}$ 和 $\{v(t),\ t\geq t_0\}$ 相关，且

$$E[w(t)v^T(s)]=S(t)\delta(t-s)\quad(t,\ s\geq t_0)$$

其中 $S(t)$ 是连续的 $p\times m$ 矩阵。

利用维纳-霍夫方程解决 $S(t)=0$ 时的最优滤波首先是由卡尔曼-布西完成的。

下面推导出来的最优滤波算法通常称为卡尔曼-布西滤波器。

我们首先分析 $t_0\leq\sigma<t$ 时，$t_1=t$ 的维纳-霍夫方程：

$$E[x(t)z^T(\sigma)]-\int_{t_0}^{t}A(t,s)E[z(s)z^T(\sigma)]ds=0$$

$$\sigma\in[t_0,\ t) \tag{4.4.21}$$

对于 $\sigma=t$ 的情形我们以后单独去分析，这样做的理由将随着一步步的分析会十分清楚的。

把式(4.4.21)对 t 求偏导，得

$$\frac{\partial}{\partial t}E[x(t)z^T(\sigma)]$$

$$= E[\dot{x}(t)z^T(\sigma)] = E\{[F(t)x(t)+G(t)w(t)]z^T(\sigma)\}$$

$$= F(t)E[x(t)z^T(\sigma)] + G(t)E[w(t)z^T(\sigma)] \qquad (4.4.22)$$

以及

$$\frac{\partial}{\partial t}\int_{t_0}^{t} A(t,s)E(z(s)z^T(\sigma))ds$$

$$= \int_{t_0}^{t} \frac{\partial A(t,s)}{\partial t} E[z(s)z^T(\sigma)]ds$$

$$+ A(t,t)E[z(t)z^T(\sigma)], \quad t_0 \leqslant \sigma < t \quad (4.4.23)$$

式(4.4.23)是根据含参量积分求导公式获得的，关于这一内容许多数学分析书中均有叙述。

现在分析式(4.4.22)的最后一项，由式(4.4.2)，有

$$z^T(\sigma) = x^T(\sigma)H^T(\sigma) + v^T(\sigma) \qquad (4.4.24)$$

于是可得

$$E[w(t)z^T(\sigma)] = E[w(t)x^T(\sigma)]H^T(\sigma) + E[w(t)v^T(\sigma)]$$

因为 $\sigma < t$ 时有 $E\{w(t)v^T(\sigma)\} = 0$，故

$$E[w(t)z^T(\sigma)] = E[w(t)x^T(\sigma)]H^T(\sigma) \qquad (4.4.25)$$

另外，方程(4.4.1)的解为

$$x(\sigma) = \Phi(\sigma, t_0)x(t_0) + \int_{t_0}^{\sigma} \Phi(\sigma, s)G(s)w(s)ds \qquad (4.4.26)$$

其中 $\Phi(\sigma, s)$ 是系统状态转移矩阵。利用这个结果有

$$E[w(t)x^T(\sigma)] = E[w(t)x^T(t_0)]\Phi^T(\sigma, t_0)$$

$$+ \int_{t_0}^{\sigma} E[w(t)w^T(s)]G^T(s)\Phi^T(\sigma, s)ds$$

$x(t_0)$ 独立于 $\{w(t), t \geqslant t_0\}$，上式右边第一项等于零.第二项中 $E[w(t)w^T(s)] = Q(t)\delta(t-s)$，$t > \sigma$ 在积分区间外，这一项也为零。因此

$$E[w(t)x^T(\sigma)] = 0, \quad t_0 \leqslant \sigma < t \qquad (4.4.27)$$

把上述结果代入式(4.4.25)，于是有

$$E[ws(t)z^T(\sigma)] = 0, \quad t_0 \leqslant \sigma < t$$

即式(4.4.22)右边最后一项为零，即

$$\frac{\partial}{\partial t} E[x(t)z^T(\sigma)]$$
$$= F(t)E[x(t)z^T(\sigma)], \quad t_0 \leqslant \sigma < t \qquad (4.4.28)$$

现在分析式(4.4.23)右边第二项，利用式(4.4.2)可知

$$E[z(t)z^T(\sigma)] = E\{[H(t)x(t) + v(t)]z^T(\sigma)\}$$
$$= H(t)E[x(t)z^T(\sigma)] + E[v(t)z^T(\sigma)] \qquad (4.4.29)$$

由式(4.4.24)，得

$$E[v(t)z^T(\sigma)] = E[v(t)x^T(\sigma)]H^T(\sigma) + E[v(t)v^T(\sigma)] \qquad (4.4.30)$$

因为 $\sigma < t$，所以 $E[v(t)v^T(\sigma)] = 0$。

进一步利用式(4.4.26)，还有

$$E[v(t)x^T(\sigma)] = E[v(t)x^T(t_0)]\Phi^T(\sigma, t_0) + \int_{t_0}^{t} E[v(t)w^T(s)]G^T(s)\Phi^T(\sigma, s)ds$$

因为 $x(t_0)$ 独立于 $\{v(t), t \geqslant t_0\}$，故上式右边第一项为零。显然

$$E[v(t)w^T(s)] = S^T(t)\delta(t - s)$$

但因 t 位于积分区间外，第二项也为零。因此

$$E[v(t)x^T(\sigma)] = 0, \quad t_0 \leqslant \sigma < t$$

结果方程(4.4.30)变为 $E[v(t)z^T(\sigma)] = 0$，$t_0 \leqslant \sigma < t$。这意味着式(4.4.29)变为

$$E[z(t)z^T(\sigma)] = H(t)E[x(t)z^T(\sigma)], \quad t_0 \leqslant \sigma < t \qquad (4.4.31)$$

于是，方程(4.4.23)可写成

$$\frac{\partial}{\partial t}\int_{t_0}^{t} A(t, s)E[z(s)z^T(\sigma)]ds$$

$$= \int_{t_0}^{t} \frac{\partial A(t, s)}{\partial t} E[z(s)z^T(\sigma)]ds$$

$$+ A(t, t)H(t)E[x(t)^T z(\sigma)], \quad t_0 \leqslant \sigma < t \quad (4.4.32)$$

合并方程(4.4.28)和(4.4.32)，就得到式(4.4.21)关于 t 的偏导数：

$$F(t)E[x(t)z^T(\sigma)] - \int_{t_0}^{t} \frac{\partial A(t,s)}{\partial t} E[z(s)z^T(\sigma)]ds$$

$$- A(t, t)H(t)E[x(t)z^T(\sigma)] = 0, \quad t_0 \leqslant \sigma < t$$

再把式(4.4.21)代入上式就得

$$[F(t) - A(t, t)H(t)] \int_{t_0}^{t} A(t,s)E[z(s)z^T(\sigma)]ds$$

$$- \int_{t_0}^{t} \frac{\partial A(t, s)}{\partial t} E[z(s)z^T(\sigma)]ds = 0 \qquad (4.4.33)$$

为使方程(4.4.33)成立，充分条件是 $A(t, s)$ 满足如下偏微分方程：

$$F(t)A(t, s) - \frac{\partial A(t, s)}{\partial t} - A(t, t)H(t)A(t, s) = 0$$

$$t_0 \leqslant s \leqslant t \qquad (4.4.34)$$

还可以证明，方程(4.4.34)是方程(4.4.33)成立的必要条件。

事实上，设

$$B(t, s) = F(t)A(t, s) - \frac{\partial A(t, s)}{\partial t}$$

$$- A(t, t)H(t)A(t, s)$$

则

$$\int_{t_0}^{t} B(t, s)E[z(s)z^T(\sigma)]d\sigma = 0$$

若 $A(t, s)$ 满足维纳-霍夫方程，则 $A(t, s) + B(t, s)$ 也满足维纳-霍夫方程。也就是说

$$\hat{x}(t|t)=\int_{t_0}^{t}A(t,s)z(s)ds$$

和

$$\hat{x}_*(t|t)=\int_{t_0}^{t}[A(t,s)+B(t,s)]z(s)ds$$
$$=\hat{x}(t|t)+\int_{t_0}^{t}B(t,s)z(s)ds$$

都是 $x(t)$ 的最优估计。由引理 4.4.1 可知，这两个估计差值的协方差矩阵必为零，即

$$E\{[\hat{x}_*(t|t)-\hat{x}(t|t)][\hat{x}_*(t|t)-\hat{x}(t|t)]^T\}=0$$

由此可得

$$\int_{t_0}^{t}\int_{t_0}^{t}B(t,s_1)E(z(s_1)z^T(s_2))B^T(t,s_2)ds_1ds_2=0 \quad (4.4.35)$$

$$\begin{aligned}E[z(s_1)z^T(s_2)]=&H(s_1)E[x(s_1)x^T(s_2)]H^T(s_2)\\&+H(s_1)E[x(s_1)v^T(s_2)]\\&+E[v(s_1)x^T(s_2)]H^T(s_2)\\&+R(s_1)\delta(s_1-s_2)\end{aligned} \quad (4.4.36)$$

由于式（4.4.36）右端头三项非负定，第四项正定[$R(s_1)$ 正定]，因此方程（4.4.35）成立的必要条件为 $B(t,s)=0$，$t_0\leqslant s\leqslant t$。

现在考察 $t_1=t$ 时的滤波方程（4.4.5），首先对它求 t 的偏导数：

$$\dot{\hat{x}}(t|t)=\int_{t_0}^{t}\frac{\partial A(t,s)}{\partial t}z(s)ds+A(t,t)z(t)$$

由方程（4.4.34）有

$$\frac{\partial A(t,s)}{\partial t}=F(t)A(t,s)-A(t,t)H(t)A(t,s)$$

于是

$$\dot{\hat{x}}(t|t)=\int_{t_0}^{t}[F(t)A(t,s)-A(t,t)H(t)A(t,s)]z(s)ds$$

$$+ A(t, t)z(t)$$

$$= A(t, t)\left[z(t) - H(t)\int_{t_0}^{t} A(t, s)z(s)ds\right]$$

$$+ F(t)\int_{t_0}^{t} A(t, s)z(s)ds$$

$$= F(t)\hat{x}(t|t) + A(t, t)[z(t) - H(t)\hat{x}(t|t)]$$

若定义 $K(t) = A(t, t)$，就得到滤波方程：

$$\dot{\hat{x}} = F(t)\hat{x} + K(t)[z(t) - H(t)\hat{x}], \quad t \geqslant t_0 \quad (4.4.37)$$

其中 $\hat{x} = \hat{x}(t|t)$，初始条件为

$$\hat{x}(t_0|t_0) = \int_{t_0}^{t_0} A(t_0, s)z(s)ds = 0$$

$K(t)$ 称为增益矩阵.

导出了式(4.4.37)的滤波方程后，我们继续推导 $n \times m$ 增益矩阵 $K(t)$ 的表达式：

现在分析 $\sigma = t$ 时的维纳-霍夫方程：

$$E[x(t)z^T(t)] - \int_{t_0}^{t} A(t, s)E[z(s)z^T(t)]ds = 0 \quad (4.4.38)$$

注意到

$$\begin{aligned} E[x(t)z^T(t)] &= E\{x(t)[H(t)x(t) + v(t)]^T\} \\ &= E[x(t)x^T(t)]H^T(t) + E[x(t)v^T(t)] \end{aligned} \quad (4.4.39)$$

$$x(t) = \Phi(t, t_0)x(t_0) + \int_{t_0}^{t} \Phi(t,s)G(s)w(s)ds$$

[由方程(4.4.26)得]

再考虑到 $x(t_0)$ 独立于 $\{v(t), t \geqslant t_0\}$，于是

$$E[x(t)v^T(t)] = \int_{t_0}^{t} \Phi(t, s)G(s)E[w(s)v^T(t)]ds \quad (4.4.40)$$

然而 $E[w(\tau)v^T(t)] = S(t)\delta(t - \tau)$，所以

$$E[x(t)v^T(t)] = G(t)S(t) \quad (4.4.41)$$

因为

$$\Phi(t,\ t)=I$$

这样，方程式 (4.4.39) 变成

$$E[x(t)z^T(t)]=E[x(t)x^T(t)]H^T(t)+G(t)S(t) \quad (4.4.42)$$

另外

$$\begin{aligned}E[z(s)z^T(t)]&=E\{z(s)[H(t)x(t)+v(t)]^T\}\\&=E[z(s)x^T(t)]H^T(t)+E[z(s)v^T(t)]\\&=E[z(s)x^T(t)]H^T(t)+E\{[H(s)x(s)\\&\quad+v(s)]v^T(t)\}\\&=E[z(s)x^T(t)]H^T(t)+H(s)E[x(s)v^T(t)]\\&\quad+R(t)\delta(t-s)\end{aligned} \quad (4.4.43)$$

把式 (4.4.42) 和 (4.4.43) 代入式 (4.4.38)，可得

$$\begin{aligned}&E[x(t)x^T(t)]H^T(t)+G(t)S(t)\\&\quad-\int_{t_0}^{t}A(t,\ s)E[z(s)x^T(t)]H^T(t)ds\\&\quad-\int_{t_0}^{t}A(t,\ s)H(s)E[x(s)v^T(t)]ds\\&\quad-\int_{t_0}^{t}A(t,\ s)R(t)\delta(t-s)ds=0\end{aligned} \quad (4.4.44)$$

由式 (4.4.26) 和 (4.4.41) 可知

$$E[x(s)v^T(t)]=\begin{cases}0, & t\neq s\\ G(t)S(t), & t=s\end{cases}$$

可见，方程 (4.4.44) 第四项被积函数在区间 $[t_0,\ t]$ 上几乎处处为零，故其积分为零。因为

$$\begin{aligned}\int_{t_0}^{t}A(t,\ s)R(t)\delta(t-s)ds&=A(t,\ t)R(t)\\&=K(t)R(t)\end{aligned}$$

方程 (4.4.44) 又可表示成

$$\begin{aligned}K(t)R(t)=&E[x(t)x^T(t)]H^T(t)\\&-\int_{t_0}^{t}A(t,s)E[z(s)x^T(t)]dsH^T(t)\end{aligned}$$

$$+ G(t)S(t)$$

$$= E\left\{\left[x(t) - \int_{t_0}^{t} A(t, s)z(s)ds\right] x^T(t)\right\} H^T(t)$$

$$+ G(t)S(t)$$

$$= E[\tilde{x}(t|t)x^T(t)]H^T(t) + G(t)S(t)$$

因为 $x(t) = \tilde{x}(t|t) + \hat{x}(t|t)$，由推论 4.4.1 可知

$$E[\tilde{x}(t|t)\hat{x}^T(t|t)] = 0$$

所以

$$\begin{aligned} E[\tilde{x}(t|t)x^T(t)] &= E\{\tilde{x}(t|t)[\tilde{x}(t|t) + \hat{x}(t|t)]^T\} \\ &= E[\tilde{x}(t|t)\tilde{x}^T(t|t)] \\ &\triangleq P(t|t) \end{aligned}$$

其中 $P(t|t)$ 为滤波误差协方差矩阵。由此可得

$$K(t)R(t) = P(t|t)H^T(t) + G(t)S(t)$$

因为 $R(t) > 0$，所以有

$$K(t) = [P(t|t)H^T(t) + G(t)S(t)]R^{-1}(t), \quad t \geqslant t_0 \tag{4.4.45}$$

现在我们导出了增益矩阵 $K(t)$ 的表达式，但又引出了滤波误差协方差矩阵。

下面再推导滤波误差协方差矩阵所满足的方程：

滤波误差及其相应的时间变化率分别为

$$\tilde{x}(t|t) = x(t) - \hat{x}(t|t) \tag{4.4.46}$$

和

$$\dot{\tilde{x}}(t|t) = \dot{x}(t) - \dot{\hat{x}}(t|t) \tag{4.4.47}$$

把式(4.4.1)，(4.4.2)及(4.4.37)代入式(4.4.47)，可得

$$\begin{aligned} \dot{\tilde{x}}(t|t) &= F(t)x(t) + G(t)w(t) - F(t)\hat{x} - K(t)[z(t) - H(t)\hat{x}] \\ &= [F(t) - K(t)H(t)]\tilde{x} + G(t)w(t) - K(t)z(t) \\ &\qquad t \geqslant t_0 \end{aligned} \tag{4.4.48}$$

由式(4.4.46)可知

$$\tilde{x}(t_0|t_0) = x(t_0) - \hat{x}(t_0|t_0) = x(t_0)$$

因此，$\tilde{x}(t_0|t_0)$ 是零均值 n 维高斯随机向量，其协方差阵为

$$\begin{aligned} P(t_0|t_0) &= E[\tilde{x}(t_0|t_0)\tilde{x}^T(t_0|t_0)] \\ &= E[x(t_0)x^T(t_0)] \\ &= P(t_0) \end{aligned}$$

因为 $\{w(t),\ t \geqslant t_0\}$ 和 $\{v(t),\ t \geqslant t_0\}$ 是零均值高斯白噪声，方程 (4.4.48) 所定义的随机过程是零均值高斯-马尔可夫过程、这样，误差过程 $\{\tilde{x}(t|t), t \geqslant t_0\}$ 完全由协方差阵所表征。

令 $\Psi(t,\ s)$ 表示系统 (4.4.48) 的状态转移矩阵，再令

$$C(t) = F(t) - K(t)H(t)$$

并把式 (4.4.48) 的解写成

$$\begin{aligned} \tilde{x}(t,\ t) &= \Psi(t,\ t_0)\tilde{x}(t_0|t_0) \\ &\quad + \int_{t_0}^{t} \Psi(t,\ s)[G(s)w(s) - K(s)v(s)]ds \end{aligned}$$

于是有

$$\begin{aligned} P(t|t) &= E[\tilde{x}(t|t)\tilde{x}^T(t|t)] \\ &= \Psi(t, t_0)E[\tilde{x}(t_0|t_0)\tilde{x}^T(t_0|t_0)]\Psi^T(t,\ t_0) \\ &\quad + E\left\{\int_{t_0}^{t} \Psi(t,\ s)[G(s)w(s) - K(s)v(s)]ds \right. \\ &\quad \left. \times \int_{t_0}^{t} [w^T(\sigma)G^T(\sigma) - v^T(\sigma)K^T(\sigma)]\Psi^T(t, \sigma)d\sigma\right\} \end{aligned} \tag{4.4.49}$$

其中 $\tilde{x}(t_0|t_0) = x(t_0)$ 独立于过程 $\{w(t),\ t \geqslant t_0\}$ 和 $\{v(t),\ t \geqslant t_0\}$，含有 $\tilde{x}(t_0|t_0)$ 和 $\{w(t),\ t \geqslant t_0\}$ 及 $\{v(t),\ t \geqslant t_0\}$ 的交叉项皆为零。因为

$$E[\tilde{x}(t_0|t_0)\tilde{x}^T(t_0|t_0)] = P(t_0)$$
$$E[w(t)w^T(s)] = Q(t)\delta(t - s)$$
$$E[v(t)v^T(s)] = R(t)\delta(t - s)$$

以及

$$E[w(t)v^T(s)] = S(t)\delta(t-s)$$

式 (4.4.49) 可以写成

$$\begin{aligned} P(t|t) = {} & \Psi(t, t_0)P(t_0)\Psi^T(t, t_0) \\ & + \int_{t_0}^{t} \Psi(t, s)\,[G(s)Q(s)G^T(s) \\ & + K(s)R(s)K^T(s) - G(s)S(s)K^T(s) \\ & - K(s)S^T(s)G^T(s)]\Psi^T(t, s)ds, \quad t \geqslant t_0 \end{aligned} \tag{4.4.50}$$

对于任意滤波器增益矩阵均可利用方程 (4.4.50) 求滤波误差协方差矩阵。但是，最优滤波增益矩阵是

$$K(s) = [P(s|s)H^T(s) + G(s)S(s)]R^{-1}(s)$$

这样，式 (4.4.50) 就成为了关于 $P(t|t)$ 的积分方程，用式 (4.4.50) 求 $P(t|t)$ 是很困难的。

为了求出 $P(t|t)$，我们把式 (4.4.50) 两边对 t 求导，可得到下面的微分方程：

$$\begin{aligned} \dot{P} = {} & \dot{\Psi}(t, t_0)P(t_0)\Psi^T(t, t_0) + \Psi(t, t_0)P(t_0)\dot{\Psi}^T(t, t_0) \\ & + \Psi(t, t)\,[G(t)Q(t)G^T(t) + K(t)R(t)K^T(t) \\ & - G(t)S(t)K^T(t) - K(t)S^T(t)G^T(t)]\Psi^T(t, t) \\ & + \int_{t_0}^{t} \dot{\Psi}(t, s)\,[G(s)Q(s)G^T(s) + K(s)R(s)K^T(s) \\ & - G(s)S(s)K^T(s) - K(s)S^T(s)G^T(s)]\Psi^T(t, s)ds \\ & + \int_{t_0}^{t} \Psi(t, s)\,[G(s)Q(s)G^T(s) + K(s)R(s)K^T(s) \\ & - G(s)S(s)K^T(s) - K(s)S^T(s)G^T(s)]\dot{\Psi}^T(t, s)ds \end{aligned} \tag{4.4.51}$$

其中，$\dot{\Psi}(t, s) = C(t)\Psi(t, s)$ $(t, s \geqslant t_0)$ 且 $\Psi(s, s) = I$ $(s \geqslant t_0)$。由式 (4.4.50) 有

$$
\begin{aligned}
\int_{t_0}^{t} \Psi(t, s)\,[G(s)Q(s)G^T(s) &+ K(s)R(s)K^T(s) \\
&- G(s)S(s)K^T(s) \\
&- K(s)S^T(s)G^T(s)]\Psi^T(t, s)ds \\
= P(t|t) &- \Psi(t, t_0)P(t_0)\Psi^T(t, t_0)
\end{aligned}
$$

把上述结果代入式(4.4.51)，于是得

$$
\begin{aligned}
\dot{P} = {} & C(t)\Psi(t, t_0)P(t_0)\Psi^T(t, t_0) \\
& + \Psi(t, t_0)P(t_0)\Psi^T(t, t_0)C^T(t) \\
& + [G(t)Q(t)G^T(t) \\
& + K(t)R(t)K^T(t) - G(t)S(t)K^T(t) \\
& - K(t)S(t)G^T(t)] + C(t)[P(t|t) \\
& - \Psi(t, t_0)P(t_0)\Psi^T(t, t_0)] + [P(t|t) \\
& - \Psi(t, t_0)P(t_0)\Psi^T(t, t_0)]C^T(t) \\
= {} & C(t)P(t|t) + P(t|t)C(t) + K(t)R(t)K^T(t) \\
& - G(t)S(t)K^T(t) - K(t)S^T(t)G^T(t) \\
& + G(t)Q(t)G^T(t), \quad t \geqslant t_0
\end{aligned}
$$

再由 $C(t) = F(t) - K(t)H(t)$，得

$$
\begin{aligned}
\dot{P} = {} & [F(t) - K(t)H(t)]P + P[F(t) - K(t)H(t)]^T \\
& + K(t)R(t)K^T(t) - G(t)S(t)K^T(t) \\
& - K(t)S^T(t)G^T(t) + G(t)Q(t)G^T(t), (t \geqslant t_0)
\end{aligned}
\tag{4.4.52}
$$

其中 $P = P(t|t)$.

方程(4.4.52)满足初始条件 $P(t_0|t_0) = P(t_0)$ 的解就是滤波方程(4.4.37)的滤波误差的协方差矩阵.

如果使用最优的增益矩阵 $K(t)$，方程(4.4.52)可以进一步简化。现在把式(4.4.52)各项重新排列写成

$$
\begin{aligned}
\dot{P} = {} & F(t)P - K(t)H(t)P + PF^T(t) - PH^T(t)K^T(t) \\
& + K(t)R(t)K^T(t) - G(t)S(t)K^T(t)
\end{aligned}
$$

$$
\begin{aligned}
&\quad - K(t)S^T(t)G^T(t) + G(t)Q(t)G^T(t) \\
&= F(t)P + PF^T(t) - K(t)[PH^T(t) + G(t)S(t)]^T \\
&\quad - [PH^T(t) + G(t)S(t)]K^T(t) + K(t)R(t)K^T(t) \\
&\quad + G(t)Q(t)G^T(t)
\end{aligned}
$$

再把 $K(t) = [PH^T(t) + G(t)S(t)]R^{-1}(t)$ 代入上式，即得到

$$
\begin{aligned}
\dot{P} = &F(t)P + PF^T(t) - [PH^T + GS]R^{-1}[PH^T \\
&+ GS]^T + G(t)Q(t)G^T(t)
\end{aligned}
$$

P 是未知项，上述方程更为方便的形式为

$$
\begin{aligned}
\dot{P} = &[F(t) - G(t)S(t)R^{-1}(t)H(t)]P + P[F(t) \\
&- G(t)S(t)R^{-1}(t)H(t)]^T - PH^T(t)R^{-1}(t)H(t)P \\
&+ G(t)[Q(t) - S(t)R^{-1}(t)S^T(t)]G^T(t) \quad (4.4.53)
\end{aligned}
$$

其中，初始条件为 $P(t_0|t_0) = P(t_0)$，$t \geqslant t_0$。

上面所有的推导可以归结为下面的定理。

定理 4.4.2 设系统可以用式 (4.4.1) 和 (4.4.2) 描述，$E[w(t)v^T(s)] = S(t)\delta(t - s)$，连续时间系统的最优滤波器由如下方程表示：

$$
\dot{\hat{x}} = F(t)\hat{x} + K(t)[z(t) - H(t)\hat{x}], \quad t \geqslant t_0 \quad (4.4.37)
$$

其中 $\hat{x} = \hat{x}(t|t)$，初始条件为 $\hat{x}(t_0|t_0) = 0$，$n \times m$ 滤波器增益矩阵为

$$
K(t) = [P(t|t)H^T(t) + G(t)S(t)]R^{-1}(t), \quad t \geqslant t_0 \quad (4.4.45)
$$

其中 $P(t|t)$ 为滤波误差 $\tilde{x}(t|t) = x(t) - \hat{x}(t|t)$ 的 $n \times n$ 协方差矩阵。

误差过程 $\{\tilde{x}(t|t),\ t \geqslant t_0\}$ 是零均值高斯-马尔可夫过程，其协方差矩阵是下列微分方程之解：

$$
\begin{aligned}
\dot{P} = &[F(t) - G(t)S(t)R^{-1}(t)H(t)]P \\
&+ P[F(t) - G(t)S(t)R^{-1}(t)H(t)]^T \\
&- PH^T(t)R^{-1}(t)H(t)P + G(t)[Q(t)
\end{aligned}
$$

$$- S(t)R^{-1}(t)S^{T}(t)]G^{T}(t) \qquad (4.4.53')$$

其中 $P = P(t|t)$ 的初始条件为

$$P(t_0|t_0) = P(t_0) = E[x(t_0)x^T(t_0)]$$

附　　录

A. 无穷维测度空间

A.1 可列个测度空间的乘积

在 3.1 节中，我们提到了无穷维乘积空间上的测度．在这里我们把有限个测度空间乘积的理论（1.10 节）推广到可列个乘积的情形。

设 $(\Omega_i, \mathscr{F}_i, \mu_i)$, $i=1, 2, \cdots$ 为概率空间序列，E 表示柱体，即

$$E = A_n \times \left(\mathop{\times}_{i=n+1}^{\infty} \Omega_i \right) \qquad (A.1.1)$$

其中 n 为自然数，而 $A_n \in \mathop{\times}\limits_{i=1}^{n} \mathscr{F}_i$。

设 $\mathscr{E}$ 是全体柱集组成的集族，容易验证 $\mathscr{E}$ 是 $\mathop{\times}\limits_{i=1}^{\infty} \Omega_i$ 中的代数。$\mathscr{F}(\mathscr{E})$ 是包含 $\mathscr{E}$ 的最小 σ 代数，$\mathscr{F}(\mathscr{E})$ 可记作

$$\mathscr{F}(\mathscr{E}) = \mathop{\times}_{i=1}^{\infty} \mathscr{F}_i \qquad (A.1.2)$$

在 $\mathscr{E}$ 上定义一个集函数：

$$\mu(E) = (\mu_1 \times \cdots \times \mu_n)(A_n) \qquad (A.1.3)$$

这里 $(\mu_1 \times \cdots \times \mu_n)(A_n)$ 表示 A_n 的有限维乘积测度。

我们可以证明(见文献[2]，p.234)：

引理 A.1.1　由式 (A.1.3) 确定的集函数 $\mu(\cdot)$ 为 $\mathscr{E}$ 上的概率测度。

定理 A.1.1　设 $(\Omega_i, \mathscr{F}_i, \mu_i)$, $i=1, 2, \cdots$ 为概率空

间序列，在 σ 代数 $\mathop{\times}\limits_{i=1}^{\infty}\mathscr{F}_i$ 上存在唯一的概率测度 μ，使得 $\mathop{\times}\limits_{i=1}^{\infty}\Omega_i$ 中的每一个柱体 $E=A_n\times\left(\mathop{\times}\limits_{i=n+1}^{\infty}\Omega_i\right)$，有

$$\mu(E)=(\mu_1\times\cdots\times\mu_n)(A_n)$$

其中 $A_n\in\mathop{\times}\limits_{i=1}^{n}\mathscr{F}_i$，$n=1,2,\cdots$。

测度 μ 称为测度 μ_i 的乘积测度，并记为 $\mathop{\times}\limits_{i=1}^{\infty}\mu_i$，于是测度空间

$$\left(\mathop{\times}\limits_{i=1}^{\infty}\Omega_i,\ \mathop{\times}\limits_{i=1}^{\infty}\mathscr{F}_i,\ \mathop{\times}\limits_{i=1}^{\infty}\mu_i\right)$$

$$\left(\mathop{\times}\limits_{i=1}^{\infty}\Omega_i,\ \mathop{\times}\limits_{i=1}^{\infty}\mathscr{F}_i,\ \mu\right)$$

称为测度空间 $(\Omega_i,\mathscr{F}_i,\mu_i)$，$i=1,2,\cdots$ 的乘积空间。

证明 因为 $\mu(\cdot)$ 是代数 $\mathscr{E}$ 上的全有限测度，根据定理 1.9.2，它可以扩张到包含 $\mathscr{E}$ 的 σ 代数上的测度。 ▮

在定理 A.1.1 中，测度空间族的指标集为全体正整数. 当指标集为任意可列集 T_c 时，只要在 T_c 与全体正整数之间建立任意的一一对应关系，定理 A.1.1 便可以推广到指标集为任意可列集的情形。

定理 A.1.2 设 (Ω_t, S_t, μ_t)，$t\in T_c$ 为一族概率空间，其中 T_c 是任意可列集，那么 σ 代数 $\mathop{\times}\limits_{t\in T_c}\mathscr{F}_t$ 上存在唯一的概率测度 μ，使得每一个柱体 $E=A_{T_n}\times\left(\mathop{\times}\limits_{t\in T_c-T_n}\Omega_t\right)$ 有

$$\mu(E)=(\mu_{t_1}\times\cdots\times\mu_{t_n})(A_n)$$

其中 $A_{T_n} \in \bigtimes_{t \in T_n} \mathscr{F}_t$，$T_n$ 是 T_c 中的有限子集，其元素为 $t_1, \cdots, t_n$。

测度 μ 称为测度 $\mu_t, t \in T_c$ 的乘积测度，并记作 $\bigtimes_{t \in T_c} \mu_t$；测度空间 $\left(\bigtimes_{t \in T_c} \Omega_t, \bigtimes_{t \in T_c} \mathscr{F}_t, \mu\right)$ 称为测度空间 $(\Omega_t, \mathscr{F}_t, \mu_t)$ 的乘积空间。

A.2 非可列无穷个测度空间的乘积

本节将测度空间乘积的理论推广到非可列无穷个因子的情形．设 T 为任意非可列无穷集，对每一个 $t \in T$，$(\Omega_t, \mathscr{F}_t, \mu_t)$ 为概率空间。

对于柱集

$$E = A_{T_c} \times \left(\bigtimes_{t \in T - T_c} \Omega_t\right) \tag{A.2.1}$$

定义集函数

$$\mu(E) = \left(\bigtimes_{t \in T_c} \mu_t\right)(A_{T_c}) \tag{A.2.2}$$

定理 A.2.1 由式 (A.2.2) 所确定的集函数 μ 是 $\bigtimes_{t \in T} \mathscr{F}_t$ 上的概率测度。

证明 显然 μ 非负，且 $\mu(\varnothing) = 0$，$\mu\left(\bigtimes_{t \in T} \Omega_t\right) = 1$。

往证 μ 具有完全可加性．设 $E_n \in \bigtimes_{t \in T} \mathscr{F}_t$，$n = 1, 2, \cdots$ 且 $\{E_n\}$ 两两不相交，不失一般性，我们可以设

$$E_n = A_{T_c}^{(n)} \times \left(\bigtimes_{t \in T - T_c} \Omega_t\right), \quad n = 1, 2, \cdots$$

其中 T_e 是 T 的可列子集，$A^{(n)}_{T_e} \in \mathop{\times}\limits_{t\in T_e} \mathscr{F}_t$，易知 $A^{(n)}_{T_e}$ 两两不相交，由定理 A.1.2 我们有

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right)=\left(\mathop{\times}_{t\in T_e}\mu_t\right)\left(\bigcup_{n=1}^{\infty} A^{(n)}_{T_e}\right)$$

$$=\sum_{n=1}^{\infty}\left(\mathop{\times}_{t\in T_e}\mu_t\right)(A^{(n)}_{T_e})$$

$$=\sum_{n=1}^{\infty}\mu(E_n)$$

所以 μ 是完全可加的。综合上面得到的结论，μ 是 $\mathop{\times}\limits_{t\in T}\mathscr{F}_t$ 上的概率测度。 ∎

测度 μ 称为测度 μ_t，$t\in T$ 的乘积，μ 可记作 $\mathop{\times}\limits_{t\in T}\mu_t$；测度空间 $\left(\mathop{\times}\limits_{t\in T}\Omega_t, \mathop{\times}\limits_{t\in T}\mathscr{F}_t, \mathop{\times}\limits_{t\in T}\mu_t\right)$ 称为测度空间 $(\Omega_t, \mathscr{F}_t, \mu_t)$，$t\in T$ 的乘积。

本书 3.1 节用到了无穷维乘积测度空间的理论。

B. 测验题答案

第一章

1. D, C, C, A, B, C, D, B, B, C.

2. (1) 由定义 $\mathscr{A}\subset\mathscr{S}(\mathscr{F})$，$\mathscr{F}\subset\mathscr{A}$. 若 $A_n\uparrow$，$A_n\in\mathscr{A}$，$\exists A_n^{(k)}\in\mathscr{F}$，$\bigcup\limits_{k=1}^{\infty}A_n^{(k)}\supset A_n$ 且 $\mu(A_n^{(k)})<\infty$，则

$$\lim_n A_n=\bigcup_{n=1}^{\infty}A_n\subset\bigcup_{n=1}^{\infty}\bigcup_{k=1}^{\infty}A_n^{(k)}$$

所以

$$\lim_n A_n \in \mathscr{A}$$

同理可证 $A_n \downarrow, \lim_n A_n \in \mathscr{A}$。因为 $\mathscr{F}$ 是代数，所以 $\mathscr{A} = \mathscr{M}(\mathscr{F}) = \mathscr{S}(\mathscr{F})$.

(2) 因为 $|f_n^2 - f^2| \leqslant 2|f_n - f|(f_n \vee f)$, $(f_n \vee f) = \sup(f_n, f)$。若$(f_n \vee f) > M$，则有 $f > \frac{1}{2}M$ 或 $|f - f_n| > \frac{1}{2}M$，否则

$$f_n = f + (f_n - f) \leqslant \frac{1}{2}M + \frac{1}{2}M = M$$

于是

$$\mu(|f_n^2 - f^2| > \varepsilon) \leqslant \mu\left(|f_n - f| > \frac{\varepsilon}{M}\right) + \mu\left(|f - f_n| > \frac{1}{2}M\right) + \mu(f > M)$$

$$\lim_n \sup \mu(|f_n^2 - f| > \varepsilon) \leqslant \mu(f > M)$$

因为 f 几乎处处有限，所以 $\lim_{M\to\infty} \mu(f > M) = 0$。

故

$$\lim_n \mu(|f_n^2 - f| > \varepsilon) = 0$$

(3) 对任意 $\varepsilon > 0$，有

$$\varepsilon^p \mu(|f_n - f| > \varepsilon) \leqslant \int_{|f_n - f| > \varepsilon} |f_n - f|^p d\mu \leqslant \int_\Omega |f_n - f|^p d\mu \xrightarrow[n\to\infty]{} 0$$

所以

$$\lim_n \mu(|f_n - f| > \varepsilon) = 0$$

第二章

1. 因为

$$E(X_{n+1}-X_n)=E[(X_{n+1}-X_n)|\mathscr{F}_n]$$
$$=E(X_{n+1}|\mathscr{F}_n)-X_n$$

所以

$$E(X_{n+1}-X_n)\leqslant 0\Longleftrightarrow E(X_{n+1}|\mathscr{F}_n)-X_n\leqslant 0$$

即

$$E(X_{n+1})\leqslant EX_n\Longleftrightarrow E(X_{n+1}|\mathscr{F}_n)\leqslant X_n,\ \forall n\geqslant 1$$
$$EX_{n+1}=EX_n\Longleftrightarrow E(X_{n+1}|\mathscr{F}_n)=X_n,\ \forall n\geqslant 1$$

2. 因为 $\{X_n\}$、$\{Y_n\}$ 一致可积，所以存在

$M>0,\ N>0$, 使 $E|X_n|\leqslant M,\ E|Y_n|\leqslant N, \forall n\geqslant 1$

故

$$E|\alpha X_n+\beta Y_n|\leqslant|\alpha|E|X_n|+|\beta|E|Y_n|$$
$$\leqslant|\alpha|\cdot M+|\beta|\cdot N,\ \forall n\geqslant 1$$

即 $\{\alpha X_n+\beta Y_n\}$ 的积分一致有界。

对任意 $\varepsilon>0$，存在 $\delta>0$，当 $P(A)<\delta$ 时，有

$$\int_A|X_n|dP<\varepsilon,\quad \int_A|Y_n|dP<\varepsilon$$

同时成立且 δ 与 n 无关。

故

$$\int_A|\alpha X_n+\beta Y_n|dP\leqslant|\alpha|\int_A|X_n|dP+|\beta|\int_A|Y_n|dP$$
$$\leqslant(|\alpha|+|\beta|)\varepsilon,\ \forall n\geqslant 1$$

即 $\{\alpha X_n+\beta Y_n\}$ 一致地绝对连续。

由定理 2.4.1 知，$\{\alpha X_n+\beta Y_n\}$ 一致可积。

3. 答：$X_{n+1}=E(X_{n+1}|\mathscr{F}_n)=X_n=\cdots=X_1$ a.s. $\forall n\geqslant$ 1。

4. 证明略。

5. $E(z_n|\mathscr{F}_{n-1})$

$$= E(X_n\chi_{\{n<\tau\}} + Y_n\chi_{\{n\geq\tau\}}|\mathscr{F}_{n-1})$$

$$= E(X_n\chi_{\{n<\tau\}} + Y_n\chi_{\{\tau=n\}} + Y_n\chi_{\{\tau\leq n-1\}}|\mathscr{F}_{n-1})$$

$$\leq E(X_n\chi_{\{\tau<n\}} + X_n\chi_{\{\tau=n\}} + Y_n\chi_{\{\tau\leq n-1\}}|\mathscr{F}_{n-1})$$

$$= E(\chi_{\{n-1<\tau\}}X_n|\mathscr{F}_n) + \chi_{\{n-1\geq\tau\}}E(Y_n|\mathscr{F}_{n-1})$$

$$= \chi_{\{n-1<\tau\}}X_{n-1} + Y_n\chi_{\{n-1\geq\tau\}} = z_{n-1}$$

又 $z_n\in\mathscr{F}_n$，故 $(z_n, \mathscr{F}_n)$ 为上鞅。

第三章

1. 由于 W_t 是布朗运动，W_t 必为鞅. 因此 $|W_t|$ 是下鞅，有

$$A|W_s| \leq E(A|W_t||\mathscr{F}_s), \forall A>0$$

e^x 是单调上升函数且 e^x 凸，利用 Jensen 不等式可得

$$e^{A|W_s|} \leq e^{E(A|W_t||\mathscr{F}_s)} \leq E(e^{A|W_t|}|\mathscr{F}_s)$$

再取数学期望，得

$$Ee^{A|W_s|} \leq Ee^{A|W_t|}, \forall t\geq s$$

于是有

$$Ee^{A|W_t|} \leq Ee^{A|W_T|} = c, \forall 0\leq t\leq T$$

2. $I_t = \int_0^t H_s dW_s = 3W_t - W_1 - W_2$

因为阶梯函数的随机积分是连续鞅，故 I_t 是连续鞅。

3. 利用伊藤公式

$$d(\sin W_t) = \cos W_t dW_t - \frac{1}{2}\sin W_t dt$$

即

$$dx_1 = x_2 dW_t - \frac{1}{2}x_1 dt$$

且

$$x_1(0) = \sin W_0 = 0$$

同理

$$d(\cos W_t) = -\sin W_t dW_t - \frac{1}{2}\cos W_t dt$$

即

$$dx_2 = -x_1 dW_t - \frac{1}{2} x_2 dt$$

且

$$x_2(0) = \cos W_0 = 1$$

4. 因为 $b(x, t) = \frac{1}{3}$，$\sigma(x, t) = x\cos t$，所以

$$|b(x, t) - b(\bar{x}, t)| = 0, |b(x, t)| < 1 + |x$$
$$|\sigma(x, t) - \sigma(\bar{x}, t)| = |x\cos t - \bar{x}\cos t| \leqslant |x - \bar{x}|$$

且

$$|\sigma(x, t)| \leqslant 1 + |x|$$

于是满足 Lipschitz 条件，故随机微分方程解存在且唯一。

5. 因为 M_t 既平稳又为鞅，故

$$E(M_t - M_s)^2 = E(E[(M_t - M_s)^2 | \mathscr{F}_s])$$
$$= E(M_t^2 - M_s^2) = r_M(0) - r_M(0) = 0$$

其中 $r_M(0) = EM_t^2 = EM_s^2$，所以

$$P\{M_t = M_0 = 常数\} = 1$$

参 考 文 献

[1] 刘文,测度论基础,辽宁教育出版社,1985.

[2] 中山大学《测度与概率基础》编写组,测度与概率基础,广东科学技术出版社,1984.

[3] 严加安,鞅与随机积分引论,上海科学技术出版社,1981.

[4] 王寿仁,概率论基础和随机过程,科学出版社,1986.

[5] A. 弗里德曼著,吴让泉译,随机微分方程及其应用(第一卷),科学出版社,1983.

[6] 陈翰馥,随机递推估计,科学出版社,1984.

[7] R. S. Liptser and A. N. Shiryayev, Statistics of Random Processes I——General Theory, Springer-Verlag, 1977.

[8] A. J. 奥斯特隆姆著,潘裕焕译,随机控制理论导论,科学出版社,1983.

[9] B. Okesendel, Stochastic Differential Equations, Springer-Verlag, 1985.

[10] R. J. Elliott, Stochastic Calculus and Application, Springer-Verlag, 1982.

[11] 复旦大学数学系编,概率论(第三册),人民教育出版社,1980.

[12] 韩崇昭、王月娟、万百五,随机系统理论,西安交通大学出版社,1987.

[13] 袁震东,自适应控制理论及其应用,华东师范大学出版社,1988.

[14] R. G. Bartle, The Elements of Integration, John Wiley & Sons, Inc., 1966.

[15] R. J. Tomkins, Martingale generalizations and preservation of martingale properties, *The Canadian Journal of Statistics*, **12** (1984), 2, 99—106.

[16] H. Robbins and D. Siegmund, A covergence theorem for non negative almost supermartingales and some applications, Optimizing Methods in Statistics, Academic Press, 1971, 233—257.

[17] Y. S. Chow, Local convergence of martingales and the law of large numbers, *The Annals of Mathematical Statistics*, **36** (1965), 552—558.

[18] A. Millet and L. Sucheston, Convergence of classes of amart indexed by directed sets, *Canadian J. Math*, **32** (1980), 86—125.

[19] M. Talagrand, Some structure results for martingales in the limit and pramarts, *Ann. of Prob.*, **13** (1985), 1192—1203.

[20] 汪振鹏,Semiamart 与 Pramart 类的收敛,华东师范大学学报(自然科学版),2(1983),27—31.

[21] 汪振鹏，鞅型序列的另一类收敛定理，应用概率统计，3(1986)，241—246.

[22] 汪振鹏，鞅型序列的局部收敛，数学年刊，**9A** (1988)，2，203—207.

[23] 汪振鹏，鞅型序列，数学物理学报，4(1985).

[24] 汪振鹏，极限鞅型序列与 GFT 收敛，数学学报，**31**(1988)，3，372—380.

[25] G. A. Edgar and L. Sucheston, Amarts: a class of asymptotic martingales——A. Discrete parameter, *J. Multivariable Analysis*, **6**(1976), 193—221.

[26] 王梓坤，随机过程论，科学出版社，1978.

[27] J. S. 梅迪奇著，赵希人译，随机最优线性估计与控制，黑龙江人民出版社，1984.

[28] H. F. Chen and L. Guo, Optimal adaptive control and consistent parameter estimates for ARMAX model with quadratic cost, *SIAM J. Control and Optimization*, **25**(1987), 4, 845—867.

[29] H. F. Chen and L. Guo, Convergence rates of least squares identification and adaptive control for stochastic systems, *Int. J. Control*, **44**(1986), 5, 1459—1476.

[30] 何声武，随机过程导论，华东师范大学出版社，1989.